



STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSO for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/986,485  
FILING DATE: 08-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/059,448  
FILING DATE: 22-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1101 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-986-485-2

Query Match 10.9%; Score 408; DB 3; Length 1101;  
Best Local Similarity 21.8%; Pred. No. 1.5e+28;  
Matches 185; Conservative 124; Mismatches 277; Indels 264; Gaps 30;  
8 LILAWV-----AGATATVPVPMWVPCPCACQIRPWYPRSSYREATTDCN 56  
18 LILWLLLEPRVPAAGPRA-----PCAACTCAGDPCACGDS-----LDGG 62  
57 DFLTAVPALPAGTQITLLQSNSTVRVDSGLYLANTE----- 97  
63 GGLALPFDLPSTWTRSLTNKLAIEDPAGFEDIPNLQEVYLNHNLTAVALSAGSS 122  
98 -----LDLSQNSFSDARDCDFHLPOLLSLHEENOL 129  
123 QVVAFLLQOQNRSLDSGLKAYLSLEVLDLNNITTEVRNTYFPHGPIKELNLAGNRI 182  
130 TRLEHSTRAGLA-SLOEYLHNHOLYRIAPRAFGSLNLLRLHNSNLLRAIDSRWEM 188  
183 GTLEGAFGDGLSRSLTLRLSKNRITQLPVRAFK-LPRLTQLDNNRRIRLIEGLTFQGL 241  
189 PNLLETIMIGGNKVDAILDMNFRPLANLRSLVLAGMNLKRISDYALEGLOSLSFYDNQ 248  
242 NSLEYLKLQRNNISKLTGAFWGLSKMHLHLEYDSLVEVNSGSLYGLTALQLHLNNS 301  
249 LARVPR-----ALEQVPGKLFPLDNKNPLOORVGPGEFAM 284  
302 IARIHRKMGSPFOKHLHELVSFNNLTRDEESLAEISSVRLSHNSISHAEGAFKGL 361  
285 LHLKELGLNNMEELVSID---KFAVLNLPETLKLIDITNNPRLSFIPRAFHLLPOMET 340  
362 RSLRYLDLDHNEISGTIEDTSCAFSGLEFGH-SKLTLEGN-KIKSVAKARAFSGLEGL 419  
341 MNNNALCALHQVIESLPNLOEVGLHGNPINCDCVIRN-----ANAT--- 383  
420 NLGVAISVDFDAFVKKNKKEHLISSDSFLCDQLKWLPPMLIGRMLOAVATCAHP 479  
384 -----GTVRFEIPSTLC-----AEP----- 400  
480 ESKQGOSIFSPPESTFVDDFLKPOITTPPTTMAVKGDIRFTGSAASSSSSPMTFAWK 539

QY 401 -----PDLR-----LPVREYF-----R 414  
DB 540 KNEVLTNADMENFVHNAQDGEVMEYITILHROVTFBHEGRYQCVITNHFSTISHA 599  
QY 415 EMTDHCPLIPSPRSPSLQVAGSESMVLHCRALAEPEEIVWTPAGLRTPAHAGRY 474  
DB 600 RLTVNLVPSFTKTPHDITIRTTVARLECAATGHPNPQIAMQKQGTDF-PAABERM 656  
QY 475 RYVP-EGTLELRVTAEEGLITCVANLVG--ADKITVSVVGRALLOPGDE-----GQ 527  
DB 657 HWPPDDVEFITDVKIDDAVYSCTAQNSAGSISANATLTLETSLVPLEDRVSVGE 716  
QY 528 GELRVQETHPYHILSWTPPNTVSTNLTWSSASSLRGOGATALARLPRTGHSYNT-- 585  
DB 717 TVALOCKATG-----NPP-----PRITW-----FKGDRLSL-----TERHHLTPD 752  
QY 586 -RL-----QATEYMACLO--VAFADHTQLACWARTREATSCHRALDRPGLAI 634  
DB 753 NQILVQNVVADAGRYCTEMSNLTGTERAHSQLSVL-----PAAGC-RKDGTVVGIPTI 806  
QY 635 LALAVLLLA 644  
DB 807 AVVSTIVLS 816

RESULT 4  
US-08-190-802A-49  
Sequence 49, Application US/08190802A  
Patent No. 5519003

GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P. O. Box 60850  
CITY: Palo Alto

STATE: CA

COUNTRY: USA  
ZIP: 94306-0850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/190, 802A  
FILING DATE: 01-FEB-1994

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:

LENGTH: 605 amino acids  
TYPE: amino acid

TOPOLOGY: unknown  
MOLECULE TYPE: protein

HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Insulin-like growth factor binding  
INDIVIDUAL ISOLATE: protein complex, Fig. 32  
US-08-190-802A-49

Query Match 10.9%; Score 407.5; DB 1; Length 605;

RESULT 6  
US-08-473-089-49  
Sequence 49, Application US/08473089  
Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Theoreof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESS: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 605 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Insulin-like growth factor binding  
INDIVIDUAL ISOLATE: protein complex, Fig. 32  
US-08-473-089-49

Query Match 10.9%; Score 407.5; DB 4; Length 605;  
Best Local Similarity 25.2%; Pred. No. 6.1e-29;  
Matches 152; Conservative 72; Mismatches 173; Indels 205; Gaps 16;

QY 3 LLLVALLAMVA-----GATATVPVPMHVPCCPQACQIRPMTTPTSSV---REATT 52  
DB 8 LALLALLSWALGPRSLGADPTGPEAGEGPACPAACV-----SYDDADELS 57  
QY 53 VDCNDLFTAVPALPAGTQTLTLLQNSIVRY-----DQSELG----- 91  
DB 58 VFCSSRNITRLPDVPGGTQALMTLDGNNISVPPAFQWLSLGLNLNLOGQLGLEPQA 117  
QY 92 ---LANLTELDLSQNSFSDARDCDFHALPOLSLHLEENQTLRL----- 133  
DB 118 LLGLENLCHLHLENRQSLALGTFAHPPALASLGLSNNRLSLRDGLGEGIGSLMDLNL 177  
QY 134 ---DHSFAGLASIQELYLHNOLYRIAPRAFSGLSMLRLHLNSNLLRAIDSRW 184  
DB 178 GWSNLAVLPDAFAGRLGSLRELVLNRLAYLPALFGLAELRELDLSRNALRAIKANV 237  
QY 185 FEMLPNLEIIMGKKNVAVI-----LDMN----- 208  
DB 238 FVQPLRLQKIVLDRLLIAVAPGAFGLKALRWLDSLNRKAVGLLEDTFPGLLGIRVRL 297  
QY 209 -----FRPLANLRSLVLAGMNLREISDYALEGLQSLSLSPYDOLARV----- 252

DB 298 SHNAIASLRPRTEKDLHLEELQLGHNRIQLAERSFEGIGOLEVTLHDNQLQEVKAGA 357  
QY 253 -----PRALEQVPGKFLDKNPQORVGGDANMHLKEL 292  
DB 358 FLGLTNVAVMNLGNCRLNLPQVFRGLKLSLHLEGSCILRIRHTTGTGSLGRPL 417  
QY 293 NNNEELVSIDKFAVLNLPETFKLDITNNPRLSFIHRAFHLPQMETLMLNNALSAL-- 350  
DB 418 KD-NGLVGLIEBSQMSIMLALLLELDLTSN-QLTHLPHRLFOGLKLEYLLLSNRRLAELPA 475  
QY 351 -----HQQTVESLPLN-----LQEVGLNG 368  
DB 476 DALGPLORAFMWLDVSHNRLEALPNSLLAPLRRLYLRLNNSLRPTTPQPGLERLMLG 535  
QY 369 NPTRDCQVIR---VANATGRV-RIIEP-----QSTLCAPPDQLRPLPR 409  
DB 536 NPMDCGCPKLALRDLALQNPASVPRFVQALCEGDCCOPPAYTYNNITCASPPVEVGLDR 595  
QY 410 EV 411  
DB 596 DL 597

RESULT 7  
US-09-063-950-2  
Sequence 2, Application US/09063950C  
Patent No. 6225085  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: MEI-019  
CURRENT APPLICATION NUMBER: US/09/063,950C  
CURRENT FILING DATE: 1998-04-21  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 673  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-063-950-2

Query Match 10.8%; Score 402; DB 4; Length 673;  
Best Local Similarity 24.9%; Pred. No. 2.3e-28;  
Matches 194; Conservative 73; Mismatches 306; Indels 206; Gaps 24;

QY 4 LVPAPLLAVAGATATVPVPMHVPCCPQACQIRPMTTPTSSSYREATTVDCNDLFLAV 63  
DB 7 LLLPILLLAG-----PEVGCPSGCGCS-----QPTVFTTARQTTV 46  
QY 64 PPALPAGTQTLTLLQNSIVRYDQSELGYLANLTELDLSQNSFSDARDCDFHALPOLSLH 123  
DB 47 PRDVPDPDVLGTVFENGITMDAGSFAGLPGLDLDLSQNLASLPSGVFQPLANLSND 106  
QY 124 LEENQTLREHDSFAGLASLOELYLHNOLYRIAPRAFSGLSNLLRLHNSNLLRAIDSR 183  
DB 107 LTANLRHEITMETFGRLERLYLGNKRIRHIOGAFDPTDLRLLELKLQDNEELALPPL 166  
QY 184 WEMLPNLEIIMGKNKYDALIDMFRPLANLRSLVLAGMNLREISDYALBELQSLSELS 243  
DB 167 ---RLPRLLDLSHNSILA-DEGILDTANVEALRLAGLQQLDEGLFSRLRLHDL 222  
QY 244 FYDNLARVPRRALEQVGLKFLDKNPQORVGGDFANMHLKELGNNNEELVSIDK 303  
DB 223 VSDNQLERV-----PVTR----- 236  
QY 304 FALVLPRLTLDITNNPRLSFIHRAFHLPQMETLMLNNALSALHQQTVESLPLNQE 363  
DB 237 ---GLRGITRLRLAGNRRIRQLRPEDLAGLAALQELDVNSISQALQGDGSLGLEPRL 292  
QY 364 VGLHGNPIRCQCVIRMAN-ATGTFVRFTIEPOSTLCAEPPDLQRLRVREVPFRFEM----- 416

QY 237 --QSLESISFYDNOQLARPRALRPOVGLKFLDNKPNLQRVGPGDPANMLHKEGL-- 292  
DB 305 LPEGIYIEIRLEONSISAKIPAGFTQYKRLKIDISKNOISDIAPDAFQGLKSLTSLVYG 364  
QY 293 NNNEELVSDIKFALVNLPELTIKDITNNPRLSFTHPRAFHHLPOMETLMMNNALSALHQ 352  
DB 365 NKITEIAKGLFDELVSF-----QILLNANKINCLRNVTFODDLONLNLSTYDKLOTISK 420  
QY 353 QYVESLPNIQEGHGNPNIRCDVIFRMANATGRVRFIEFQSTLCAEPDLORLPVREV- 411  
DB 421 GLFAPLQSIOTLTLAONPFVCDCHEKWL-ADYLQDNPIETSGARCSPPRLANKRISQIK 479  
QY 412 --PFR-----EMTDHC-LPLISPR-----SFPSSLOY 435  
DB 480 SKKFRSGSSEDIYRSRFSSECFMDLVCPEKRCBGITVDCSNOKLVRIPIPHLPEYVDLRL 539  
QY 436 ASGESWYLHCRALAEPEPEIYWTTPA-----GLTLTAHAG 471  
DB 540 NDNEVSVLATGTGFKLPNLRKINLSNNKIKVEGAFDGAASVQELMLTGNLETVH-G 598  
QY 472 RRYR-VYPEGTLELRVTAEAGLYTCAONLV-----GADTKT 509  
DB 599 RYRGLSGIKTLMR-----SLLISCVSNDFAGLSVRLSLYDNRTITTPGAFTTL 652  
QY 510 VSVYVGRALLQGRDRGQGLELRVQETHPHYHILLSVTPPTVSTN 555  
DB 653 VSPVHKKPEVOP-----LOLQL-----PLAWLGKWLKRRRIYSGN 687

## RESULT 10

US-08-190-802a-50  
Sequence 50, Application US/08190802a  
Patent No. 5519003

## GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190,802a

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 603 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Insulin-like growth factor bind.

INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33

US-08-190-802a-50

Query Match 9.2%; Score 344; DB 1; Length 603;

Best Local Similarity 23.3%; Pred. No. 4, 5e-23;

Matches 141; Conservative 63; Mismatches 194; Indels 208; Gaps 16;

QY 8 LLLAWVA-----GATATPVVPMHVPCPCACQIRMYVPRSSYREATTVDGN 56  
DB 13 LLAFWALGPCHLQGDPPGASDAE-----GPCPVACTS-----HDDYDELVSFCS 61  
QY 57 DFLTAVPALPAGTQTLQNSIVRV-----DOSELGY 91  
DB 62 SKNLTLPPDIPYSTRALMDGNLSSIPSAFQNLSDFLNLQGSWLRSLRQALLG- 120  
QY 92 LANLFLDLSQNSFSDARCDHALPQLSLHLEENQTLRE----- 133  
DB 121 LQNLVYLHLERNRLMAVGLFTHTPSLASSLSSMLGLREGLRQGLSHLMDLNGWN 180  
QY 134 -----DHSFAGLASQELYNHNLRYAPRAFSGLSMLRLNLSNLLRAIDSRPFEM 187  
DB 181 SLVLPDYVFOGLGNLHELVLGKNLTYLQPALFCGLGELREILUSRNALRSYKANVEVH 240  
QY 188 LPNLETIMIGNKVDAL-----LDNN----- 208  
DB 241 LPRLOKLYIDRNLTAVAGAFGLMKALRWLDLSHNRVAGLMEDTPRGLGLHYLRALHN 300  
QY 209 -----FRPLANRSLVLAGMNLREISDYALBGLQSLSESLSFYDNOARV----- 252  
DB 301 AIALSRPRTFKDLHLEELQGLHNRIRQGERFEGLEGOLEVTLNDNGITEVRVAGAFSG 360  
QY 253 -----PRALRQVPGKFLDNKPNLQRVGPGDPANMLHKEGLNMM 295  
DB 361 LFNVAVMNLSGNCLSLPRFVFOGLKLSLHLEHSCIGHVHLHTFAGLSGRRLFLRD- 419  
QY 296 EELVSDIKFALVNLPELTIKDITNNPRLSFTHPRAFHHLPOMETLMMNNALSAL----- 350  
DB 420 NSISIEBQSLAGLSLELBDLTN-RLTHLRLQPLFOGLGHEYLILSLNQLTTLTSAEVL 478  
QY 351 -----HQ-----QVESLPNIQEGHGNPNIR 372  
DB 479 GPLORAFWIDISHNLETLAEGLSLGRVYLSLRNNLSQFFSPQGLERLMDLNDPMD 538  
QY 373 CDCVIR-----WANATGTVRFIEP-----QSTICARPDQRLPVREVPFR 414  
DB 539 CSCPLALRDFALQNPVGVPFVQVCEGDCQPVYTYNNITCAGAPANVSGIDLNDVSET 598  
QY 415 EMTDHC 420  
DB 599 HPV-HC 603

## RESULT 11

US-08-477-346-50  
Sequence 50, Application US/08477346  
Patent No. 6262023

## GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

Db 241 LPRLOKLYIDRNLITRVAAGATLGMKALRWLDLSHNRVAGLMEDETPGGLGLHVLRLAHN 300  
OY 209 -----FRPLANRLSVLAGMNLREISDYALEGLOSLESFYNOLARV----- 252  
Db 301 ATASLRPRFKLHLEELQLOHNRIROIGERTFEGLOLEVLITLNDNOITTEVRVAESG 360  
OY 253 -----PRALAEVPGKFLDLINKNPLORVGPGDEANMLHLKELGLANM 295  
Db 361 LFNVAVMNLGNCRLNSRPRVFOGLDKLSLHLEHSCIGHVRLHTEFAGLSGLRRLFLRD- 419  
OY 296 BELVSIDKFAVLNPELTKLDITNNPRLSFIRPRAFHLPQMETMLNNAISAL----- 350  
Db 420 NSISIEROSLAGLSLELDELITTN-RLTHPLROLFOGLGHEYLILLSYNOLTTLSAEVL 478  
OY 351 -----HQ-----OTVESLPNLOEVGLHGNPRL 372  
Db 479 GPLORAFMLDISHNLETLAEGLFSSLGVRVYLSLRNNSLOTFSPQPLERLMLDANPMD 538  
OY 373 CDCVIR-----WANATGTRVRPIEP-----OSTLCAEPPDLORLPVREVPFR 414  
Db 539 CSCPLKALRDFALONPVPVRFVOTVCBGDDCQPYVYTNITCAGAPANVSGLDLRVSFT 598  
OY 415 EMTDHC 420  
Db 599 HFV-HC 603

RESULT 13  
US-09-191-647-2  
Sequence 2, Application US/09191647  
Patent No. 6046015  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey  
APPLICANT: Kid, Thomas  
APPLICANT: Brose, Katja  
APPLICANT: Tessier-Lavigne, Marc  
TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
FILE REFERENCE: B98-031-3  
CURRENT APPLICATION NUMBER: US/09/191,647  
CURRENT FILING DATE: 1998-11-13  
EARLIER APPLICATION NUMBER: 60/065,544  
EARLIER FILING DATE: 1997-11-14  
EARLIER APPLICATION NUMBER: 60/081,057  
EARLIER FILING DATE: 1998-04-07  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1525  
TYPE: PRT  
ORGANISM: human  
US-09-191-647-2

Query Match 8.9%; Score 331; DB 3; Length 1525;  
Best Local Similarity 17.8%; Pred. No. 3.4e-21;  
Matches 126; Conservative 76; Mismatches 177; Indels 330; Gaps 11;

OY 29 CPQCACQIRPWYTPRSSYREATVDCNDLFLTAVPALPAGOTOTLLQNSIYRVDOSE 88  
Db 28 CPAQCSC-----SGSTVDCGHLALRSVPNRPNTERTLIDNGNNITRITTD 74  
OY 89 LGYLANITELDLSONSFSDARDCDFHALPOLLSLHLENOQLTRLEDHSFAGLASIQEYL 148  
Db 75 FAGLRHLRVQLMKNKISTIERGAFODLKELERLNLNNHLOLPELLEFLGTAKIYRDL 134  
OY 149 NHNOLYRIAPRAFSGLSMLRLHLNSNLLRAIDSRWFMPLNLEITIMIGKKVDAI----- 204  
Db 135 SENIOIAPRAFGAVIDKMLDLYNOISCIEDGAFRALRDLLEVLTINNNTIRLSVAS 194  
OY 205 ----- 204  
Db 195 FNHMPKILTRFLHSHNNLYCCHLAWLSDMLAKKRPVGLTYTQCMGSHLRGNVAEVRRE 254

OY 205 -----LDNM 208  
Db 255 FVCSDEEBHQSMAPSCSYLHCPAACTCSNNIVDCRGKLEIFPNLPETTEIRLEON 314  
OY 209 -----FRPLANRLSVLAGMNLREISDYALEGLOSLESFYNOLARVRR----- 255  
Db 315 TIKVIPGAFSPYKKRRLRDLSSNQLSELAPPAFQGLRSLNSLYLGNKITEELPKSLFEG 374  
OY 256 -----ALEVPGKFLDLINKNPLORVGPGDEANM----- 284  
Db 375 LPSLOLLLNANKINCLRYDAFODLHNLMLSLYDKRLQTIAGFTSPRLAIOYTHLAQN 434  
OY 285 -----LHUK----- 288  
Db 435 PRICDCHLWLDYHTNDPIETSGARCTSPRLANKRIGQISKKRCSTGEDYRSKLSG 494  
OY 289 -----ELGINNMEELYSIDKFAVLNLP 310  
Db 495 DCFADLACEKRCRGESTVDCSNQKNTIPEHITPOYTABLRNLNNEFTVLEATGIFKLP 554  
OY 311 ELTKLDITNP-----RLSFTHPRAFHLPQMETMLNNAI 347  
Db 555 QLRKINFNNKTTDIEGAFEGASGVNEILTSNRLENVOHKMFKGLJESLKTMLRSNRI 614  
OY 348 SALHQQT-----VESLPNLOEVGLHGNPRLCDCVIRANAT 383  
Db 615 TCVGNDSEFIGLSSVRLSLYDNOITTVAPGADTSLHSLTLNLANPFCNOCYLAWLGM 674  
OY 384 GTRVRFI--EPOSTLCAEPPDLORLPVREVPREMT-----DHCLPL 423  
Db 675 LKKRIVGNPR---CQRYPLEKEIPIDVALQDTCTCDGNDNDCSPL 720

RESULT 14  
US-09-540-245A-2  
Sequence 2, Application US/09540245A  
Patent No. 6270984  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey  
APPLICANT: Kid, Thomas  
APPLICANT: Brose, Katja  
APPLICANT: Tessier-Lavigne, Marc  
TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
FILE REFERENCE: B98-031-3  
CURRENT APPLICATION NUMBER: US/09/540,245A  
CURRENT FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 60/065,544  
PRIOR FILING DATE: 1997-11-14  
PRIOR APPLICATION NUMBER: 60/081,057  
PRIOR FILING DATE: 1998-04-07  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1525  
TYPE: PRT  
ORGANISM: human  
US-09-540-245A-2

Query Match 8.9%; Score 331; DB 4; Length 1525;  
Best Local Similarity 17.8%; Pred. No. 3.4e-21;  
Matches 126; Conservative 76; Mismatches 177; Indels 330; Gaps 11;

OY 29 CPQCACQIRPWYTPRSSYREATVDCNDLFLTAVPALPAGOTOTLLQNSIYRVDOSE 88  
Db 28 CPAQCSC-----SGSTVDCGHLALRSVPNRPNTERTLIDNGNNITRITTD 74  
OY 89 LGYLANITELDLSONSFSDARDCDFHALPOLLSLHLENOQLTRLEDHSFAGLASIQEYL 148  
Db 75 FAGLRHLRVQLMKNKISTIERGAFODLKELERLNLNNHLOLPELLEFLGTAKIYRDL 134  
OY 149 NHNOLYRIAPRAFSGLSMLRLHLNSNLLRAIDSRWFMPLNLEITIMIGKKVDAI----- 204

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:34:58 ; Search time 35.32 Seconds

(without alignments)  
2242.231 Million cell updates/sec

Title: US-09-905-088a-245

Perfect score: 3732  
Sequence: 1 MRLIVAPLLAMVAGATATV.....RKLPSSGEETILPPLSQNS 713

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3732	100.0	713	20	AAV13385
2	3732	100.0	713	22	AAB80253
3	1993	53.4	377	22	AAB92765
4	1695	45.4	716	21	AAB33472
5	1695	45.4	716	22	AAU12427
6	1695	45.4	716	22	AAB27234
7	1695	45.4	716	22	AAB50965
8	1695	45.4	716	22	AAG67530
9	1672	44.8	716	22	AAAM0376
10	1656.5	44.4	705	21	AAB42513
11	1656.5	44.4	705	22	AAAM78823

12	1656.5	44.4	708	20	AAV13355	Amino acid sequenc
13	1656.5	44.4	708	21	AAV70030	Human extracellular
14	1656.5	44.4	708	22	AAAM39309	Human polypeptide
15	1656.5	44.4	708	22	AAU12320	Human PRO220 polyp
16	1656.5	44.4	708	22	AAB80223	Human PRO220 prote
17	1654.5	44.3	719	22	AAV79807	Human protein SEQ
18	1654.5	44.3	719	22	AAAM4095	Human polypeptide
19	1640	43.9	707	20	AAV02379	Polypeptide identi
20	1504	40.3	292	22	AAE03524	Human secreted pro
21	875.5	23.5	431	22	AAE94521	Human protein sequ
22	753.5	20.2	273	22	AAE93523	Human protein sequ
23	546.5	14.6	592	22	AAE09437	Human sbgtrango79a
24	492.5	13.2	620	22	AAE07405	Human membrane ass
25	490.5	13.1	614	20	AAW84596	Amino acid sequenc
26	489.5	13.1	620	20	AAV13357	Amino acid sequenc
27	489.5	13.1	620	22	AAU12333	Human PRO227 polyp
28	489.5	13.1	620	22	AAE08025	Human PRO227 prote
29	468.5	12.6	606	22	AAE06799	Human neuronal gut
30	467.5	12.5	579	22	AAE06804	Mature human neuro
31	442	11.8	1091	18	AAW41641	Sequence used in d
32	442	11.8	1091	20	AAV08099	Murine glial cell
33	442	11.8	1091	20	AAV08010	Mouse LIG-1 protei
34	442	11.8	1091	21	AAV97833	Murine LIG-1 prote
35	438.5	11.7	640	20	AAV08100	Human PRO331 prote
36	438.5	11.7	640	20	AAW85722	Novel protein (Cio
37	438.5	11.7	640	20	AAV13394	Amino acid sequenc
38	438.5	11.7	640	21	AAE24407	Human PRO331 prote
39	438.5	11.7	640	21	AAV70673	Human PRO331 prote
40	438.5	11.7	640	22	AAU12355	Human PRO331 polyp
41	438.5	11.7	640	22	AAU00826	Human immune respo
42	438.5	11.7	640	22	AAE08026	Human PRO331 prote
43	438.5	11.7	640	22	AAE65292	Human angiotensin
44	438.5	11.7	640	22	AAE53089	Human angiotensin
45	432.5	11.6	540	22	AAE65612	Drosophila melanog

#### ALIGNMENTS

RESULT 1	AAV13385	standard; Protein; 713 AA.
ID	AAV13385;	
XX	AAV13385;	
XX		
DT	25-JUN-1999	(first entry)
XX		
DE	Amino acid sequence of protein PRO293.	
KW	Secreted protein; transmembrane protein; human; enterocolitis;	
KW	Zollinger-Elison syndrome; gastrointestinal ulceration;	
KW	congenital microvillus atrophy; skin disease; cell growth;	
KW	abnormal keratinocyte differentiation; porriasis; epithelial cancer;	
KW	Parkinson's disease; Alzheimer's disease; ALS; neuropathy;	
KW	fibromodulin; dermal scarring; Usher Syndrome; Atrophila areata;	
KW	anti-thrombotic; wound healing; tissue repair.	
OS	Homo sapiens.	
XX		
PN	WO9914328-A2.	
XX		
PD	25-MAR-1999.	
XX		
PF	16-SEP-1998;	98MO-US19330.
XX		
PR	25-NOV-1997;	97US-0066840.
PR	17-SEP-1997;	97US-0059113.
PR	17-SEP-1997;	97US-0059115.
PR	17-SEP-1997;	97US-0059117.
PR	17-SEP-1997;	97US-0059119.
PR	17-SEP-1997;	97US-0059121.
PR	17-SEP-1997;	97US-0059122.
PR	17-SEP-1997;	97US-0059184.

KW antiarthritic; antifertility; antidiabetic; antiviral; diabetes;  
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;  
KW ischaemia; inflammation.  
XX  
OS Homo sapiens.  
PN MO200104311-A1.  
XX 18-JAN-2001.  
PD  
XX 22-FEB-2000; 2000MO-US04414.  
PF  
XX  
PR 07-JUL-1999; 99US-0143048.  
PR 26-JUL-1999; 99US-0143698.  
PR 28-JUL-1999; 99US-0146222.  
PR 08-SEP-1999; 99MO-US20594.  
PR 13-SEP-1999; 99MO-US20594.  
PR 15-SEP-1999; 99MO-US21090.  
PR 15-SEP-1999; 99MO-US21547.  
PR 05-OCT-1999; 99MO-US23089.  
PR 29-NOV-1999; 99MO-US28214.  
PR 30-NOV-1999; 99MO-US28313.  
PR 16-DEC-1999; 99MO-US30095.  
PR 20-DEC-1999; 99MO-US30911.  
PR 20-DEC-1999; 99MO-US30999.  
PR 05-JAN-2000; 99MO-US00219.  
XX  
XX (GENTH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;  
PI Filvaroff E, Fong W, Geiber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kijavini JF;  
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;  
PI Williams PM, Wood WI;  
XX  
DR WPI: 2001-081051/09.  
DR N-PsDB: AAF72414.  
XX  
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in  
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung  
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.  
PT Alzheimer's disease) -  
XX  
XX  
PS Claim 1; Fig 86; 393pp; English.  
XX  
CC The present sequence is one of sixty one novel secreted and  
CC transmembrane PRO polypeptides. The PRO polypeptides are  
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung  
CC squamous cell carcinoma), gastrointestinal disorders (e.g.  
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.  
CC endometrial bleeding angiogenesis, ischaemias such as coronary  
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,  
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and  
CC diabetes and retinal disorders such as retinitis pigmentum.  
CC The PRO nucleic acids have applications in molecular biology, including  
CC use as hybridization probes, and in chromosome and gene mapping.  
XX  
SQ Sequence 713 AA:  
  
Query Match 100.0%; Score 3732; DB 22; Length 713;  
Best Local Similarity 100.0%; Pred. No. 1e-290;  
Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLLVAPLLAVVAGATATVPVPMVHPCPCACQIRPWYPRSSRYRATVDCNDL 60  
DB 1 MLLVAPLLAVVAGATATVPVPMVHPCPCACQIRPWYPRSSRYRATVDCNDL 60  
QY 61 TAVPALPAGTGTLLTQNSIVRVDSGLYLANLELDSNFSFSDARDCEHALPOL 120  
DB 61 TAVPALPAGTGTLLTQNSIVRVDSGLYLANLELDSNFSFSDARDCEHALPOL 120

QY 121 SLHLEENQTRLEDHSPFAGLASIQELYLNHNQYRIAPRAFSGSLNLRNLRAI 180  
DB 121 SLHLEENQTRLEDHSPFAGLASIQELYLNHNQYRIAPRAFSGSLNLRNLRAI 180  
QY 181 DSRWFEMLPNLEIIMIGKVFDAIILDMNRPPLANLSVLAGMRLREISYALEGLOSLE 240  
DB 181 DSRWFEMLPNLEIIMIGKVFDAIILDMNRPPLANLSVLAGMRLREISYALEGLOSLE 240  
QY 241 SLSPFYDNQALVRPRALAEQVPGIKFLDLNKNPQLQVPGSGFANLHLKEGLNMELYS 300  
DB 241 SLSPFYDNQALVRPRALAEQVPGIKFLDLNKNPQLQVPGSGFANLHLKEGLNMELYS 300  
QY 301 IDKFAVLNPELTKLDTNNPRLSEFIHPRAFHMLPQMETLNNNALSAALHOOTVESLPN 360  
DB 301 IDKFAVLNPELTKLDTNNPRLSEFIHPRAFHMLPQMETLNNNALSAALHOOTVESLPN 360  
QY 361 LQEVGLHGNPFRCDVYIRMANATGTRVRETEPOSTLCAEPPDLQRLPVRVPREMTDHC 420  
DB 361 LQEVGLHGNPFRCDVYIRMANATGTRVRETEPOSTLCAEPPDLQRLPVRVPREMTDHC 420  
QY 421 LPLISPRSPPSIOVAGSESMVHCRALAPPEPEITWVTPAGLRLPFAHGRXYRYPG 480  
DB 421 LPLISPRSPPSIOVAGSESMVHCRALAPPEPEITWVTPAGLRLPFAHGRXYRYPG 480  
QY 481 TLELRRTAEAGLYTCVAONLGVADTKTVSVYVGRALDPGDEGGLELRVOETHPYH 540  
DB 481 TLELRRTAEAGLYTCVAONLGVADTKTVSVYVGRALDPGDEGGLELRVOETHPYH 540  
QY 541 ILLSWTPPTVSTNLTWSSASSLRGOGATALALRPGTHSYNITRLLQATEWYACLOVA 600  
DB 541 ILLSWTPPTVSTNLTWSSASSLRGOGATALALRPGTHSYNITRLLQATEWYACLOVA 600  
QY 601 FADAHQOLACVMARTKEATGCHRALDRGLAIILAAVILAAAGLAHGTGQPRKGVG 660  
DB 601 FADAHQOLACVMARTKEATGCHRALDRGLAIILAAVILAAAGLAHGTGQPRKGVG 660  
QY 661 GRRLPPLPAMFQMSAPSVRVVSAPLVPNPGKILPRSSSEGETLPLPSSONS 713  
DB 661 GRRLPPLPAMFQMSAPSVRVVSAPLVPNPGKILPRSSSEGETLPLPSSONS 713  
  
RESULT 3  
ID AAB92765  
ID AAB92765 standard; Protein; 377 AA.  
XX  
AC AAB92765;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:11234.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI: 2001-318749/34.





```

Db      210  nfkriplandrsilvagmyttdipgnalvgldeslsfydnklvkvnpqalqkvpnlfkd 2656
QY      268  LNKNPDLQRYGPDGFANMLHLKEGLINNMDELVSIDKFLALVLPETLTKLDITNNRPLSFII 327
Db      270  lnknpihkiqegdfkxmrlrklvelglnmmgelvsvdryaldnlpelctkreatmpkisyih 3299
QY      328  PRAFIHLQMETLMLNNNALSAHQOYESLPNLQEWGLHGPNIRCDCVIRMANATGRV 3877
Db      330  rlafrsvpaleslmnnalnaiyuklvesipnlreishnprlcvdcvlihwlnskotnl 3899
QY      388  RTIEQSTLCAEPRLDQLPVREVPFRKMTDHCPLSPSPSPSLQVASGESWYVHCRA 4477
Db      390  rfmeplsmfcampeykyhqnkeyvlqdsseqclpmshdsfpnrlnvdigtvtflcdra 4499
QY      448  LAEPRELYWYTPRGLRTPLRHACRGRRVYDEGTELEIRNTYAEEAGLYTCVAQNLVADRT 5077
Db      450  maepereplywtrpynkltvetclsdckykliseegleltsnqldeossyrttcvaqnvpgadt 5099
QY      508  KTVSVVVGRRALLQGRDEGQGLERVRQDETNYHILLSWMPRPNTVSNLWSSAS-SLRG 5666
Db      510  rvaalkvngtll-----dgtqylaklyvngqteshllsvakxnsnwmvstnllwssaatmkidn 5655
QY      567  QGATRALARLPRTGHSYNTIRLLQATQYEWACIAQVAFADAHQLOACVMARTEKA----- 6118
Db      566  phlctayrardvdehynulthipqetdyevcltvsnlnqgqkscvovntknaafavdisd 6255
QY      619  -----TSCRHALGDRPGLAILAILALV 639
Db      626  getataalaavmgsmfavilasiaslav 650

```

RESULT	7
AAB50965	
ID	AAB50965 standard; Protein; 716 AA.
XX	
AC	AAB50965;
XX	
DT	21-MAR-2001 (first entry)
XX	
DE	Human PRO1338 protein.
XX	
KW	Human; PRO; cytostatic; neuroprotective; respiratory general; antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant; PMO agonist; cancer; inflammatory disorder; Immunological disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200073348-A2.
XX	
PD	07-DEC-2000.
XX	
PF	30-MAY-2000; 2000WO-US14941.
XX	
02-JUN-1999;	99WO-US12252.
PR	22-JUN-1999; 99US-0140650.
PR	23-JUN-1999; 99US-0141037.
PR	20-JUL-1999; 99US-0144758.
PR	01-SEP-1999; 99WO-US20111.
PR	08-SEP-1999; 99WO-US20594.
PR	29-OCT-1999; 99US-0162506.
PR	30-NOV-1999; 99WO-US28313.
PR	01-DEC-1999; 99WO-US28634.
PR	02-DEC-1999; 99WO-US28551.
PR	16-DEC-1999; 99WO-US30095.
PR	20-DEC-1999; 99WO-US30999.
PR	06-JAN-2000; 2000WO-US00376.
PR	11-FEB-2000; 2000WO-US03565.
PR	18-FEB-2000; 2000WO-US04341.
PR	18-FEB-2000; 2000WO-US04342.
PR	02-MAR-2000; 2000WO-US05841.
PR	03-MAR-2000; 2000US-0187202.
PR	10-MAR-2000; 2000WO-US06319.

PR	15-MAR-2000; 2000OWO-US06884.
PR	30-MAR-2000; 2000OWO-US08439.
PR	17-MAY-2000; 2000OWO-US13705.
XX	
XX	
PA	(GETH ) GENENTECH INC.
PI	
PI	Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC,
PI	Shelton DL, Smith V, Watanabe CK, Wood WI;
DR	WPI: 2001-016509/02.
XX	N-PSDB; AWC91567.
XX	
PT	Twenty eight nucleic acids encoding PRO polypeptides which are useful
PT	for treating various tumors, e.g. breast cancer, and other
PT	inflammatory, angiogenic and immunological disorders -
XX	
PS	Claim 31; Fig 30; 188pp; English.
XX	
CC	The present sequence is one of twenty eight novel PRO polypeptides. The
CC	PRO polypeptides and their agonists, including antibodies, peptides, and
CC	small molecule agonists, may be used to treat various tumors, e.g.,
CC	cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
CC	cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
CC	central nervous system cancer, melanoma or leukemia. They are also
CC	useful for treating other disorders such as neuronal, glial, astrocytal,
CC	hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC	biological disorders, and inflammatory, angiogenic and immunological
CC	disorders.
XX	
XX	Sequence 716 AA;
50	

Query Match	45.4%	Score 1695	DB 22	Length 716
Best Local Similarity	53.6%	Pred. No. 3.5e-127		
Matches 333	Conservative 104	Mismatches 166	Indels 20	Gaps 5
QY	CPPOCACIRWYWRSSYREATYVDCNDLEFATVPALPACTOFTLLLOSNIVR-VDOS	87		
DB	32 cpqlcveclrpwfpqstgyreatvtrdcndrlrttrpnsalsctyqlllqsmniaktvd--	89		
QY	88 ELGYANTLEITDLSQNSFSARDCDFHALPOLLSHLEENOLTILEDHSFAGLSLOELY	147		
DB	90 elqqlfnlteldfsqnfntikexyglanllqtlthleengttemtdyqlqdlsmngely	149		
QY	148 LNHNOYLIAARRAAGSLNLRHNSYLRAIDSRWEMLPNLFILMIGCKYDALIDM	207		
DB	150 lnnhgslstshatagllknllrtlnhnskllvldstwtidscpnlleilmagenprvgllidm	209		
QY	208 NEFRLANLRSVLAGMNLREISDALSEQLSESLSPFDNOLARVPRALDEQVPELTKFD	267		
DB	210 nfkrlanlrstslagmlyltdlpnmalvgldsestsfydnklvkrpqalqakvpnlkfld	269		
QY	268 LNKNPLORVGCDPRANMLHLKELGLNNKEBELVSDIKFALVNLPELTKIDITNNPRLSFIH	327		
DB	270 lnnkpnlnkigeddkmmrlrkeldginnmgelvsvdygaldlpeitkileacnnpklsyih	329		
QY	328 PRANHLPOWMTLLNNNALSGLHOQFYESLNIOEVLNHNPIRCCOVIMANATGRV	387		
DB	330 rlatrsypalteslmmnalaiyqklvesipnlrtreislnhprltdcvtlmlwnsknlti	389		
QY	388 RFIEPOSTLCAEPDLOLRPLREVPFRFEMTDHCLPLISPRSFPELSLOVAGSESMWLHCRA	447		
DB	390 rfimeprlsmfcmpeygyhkeyvlidssseqclpmshdsfipnrlnvdlrtvflidcra	449		
QY	448 LAEPEPELTYWTPRAGLNLPHNAAGRRKRVREEGTLELRRTVAEBAGLYTCVAQNLVGADT	507		
DB	450 maepepeiywtpriqnlklyetelsckyklsesgflseisnldiedsgrytcvaqnvgsdct	509		
QY	508 KTVSVWVGRALOGRGDGGGLERVOETHRYHLLSWVPNPNVSNLTMSSAS-SLRG	566		
DB	510 rvatikxngcll-----dgtvklkyvqcheesllvswkxnsnmctsnllkxssatmkldn	565		
QY	567 OGATATLAREPGTHSYNITRLLQATEEWACLOVFAADANTOLACVMAATKRA-----	618		

PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000MO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
PA  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR N-PSDB; AA159532.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries.  
XX  
PS Example 6; SEQ ID NO 3521; 10078bp; English.  
XX  
CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA38642-AA42213) with nucleotide, and  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 716 AA:  
SQ

Db 330 rlafrswpaleslmnnalnnaiaiyqktvesjpnltreisibnsprlrcdcovlhrinsnktui 389  
Qy 388 RFRPQSTLCAEPDRLRVPYREVPFREMDDHCPPLSPSPFPISLOVASESNVHLHORA 447  
Db 390 rfmepismfcmpeykyghvkeylqdsseqclpmshsfipnrlnvdlgtvflclora 449  
Qy 448 LAEPEPEIYWTTPAGLRTPAHAGRRYRVYDEGFLTEIRRYTAERAGLYTCVAOVLVADDT 507  
Db 450 maepepeiyywtvtpignkivtelcldskylsseglteislndqledsgrtvcvaqnvqgadt 509  
Qy 508 KTVSVVVGRRALPGRDEGGLERVOETHPYHLLSWTPPMTVSTNLMSSAS-SLRG 566  
Db 510 rvalikvngtll-----dgtvklkyvqtshsalvskvsnvmtelnkwssetmkidn 565  
Qy 567 QGATALARLRGTHSYNTITPLQATEYACLOVAFADAHOTLACWARTREA----- 618  
Db 566 phlytarvpvdvheynlthlpdstdyevcltvsnlhngqkscvntknaafavdisd 625  
Qy 619 ----TSCRRALGDRPGILATLALAV 639  
Db 626 getstalaavmgsmfavisiaslav 650  
RESULT 10  
AAB42513  
ID AAB42513 standard; Protein; 705 AA.  
AC AAB42513;  
DT 08-FEB-2001 (first entry)  
XX Human ORFX ORF2277 polypeptide sequence SEQ ID NO:4554.  
XX  
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnerrary; antiparkinsonian; antiparkinsonian; nucleotropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antihypoid;  
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.  
XX  
KW Homo sapiens.  
OS  
XX  
XX WO200056473-A2.  
XX  
XX 05-OCT-2000.  
XX  
XX 31-MAR-2000; 2000MO-US08621.  
XX  
XX 31-MAR-1999; 990US-0127607.  
XX 02-APR-1999; 990US-0127636.  
XX 05-APR-1999; 990US-0127728.  
XX 30-MAR-2000; 2000US-0540763.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach M;  
XX  
XX WPI; 2000-602362/57.  
XX N-PSDB; AAC76722.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
XX  
XX Claim 11; Page 3752-3754; 5507bp; English.  
PS







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:34:58 ; Search time 24.35 Seconds

(without alignments)  
2813.622 Million cell updates/sec

Title: US-09-905-088a-245

Perfect score: 3732

Sequence: 1 MRLVAPLLAMVAGATATV.....KKLPSSGEGTLLPPLSONS 713

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Database : PIR\_71:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

#### SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1649.5	44.2	707	2 JCT763	neuronal leucine-r
2	442	11.8	1091	2 A58532	gliad cell membran
3	407.5	10.9	605	2 A41915	insulin-like growt
4	397.5	10.7	605	2 JCT5239	insulin-like growt
5	381.5	10.2	789	2 T28714	hypothetical prote
6	381.5	10.2	1355	2 T28715	hypothetical prote
7	368	9.9	907	2 JG0193	G protein-coupled
8	360.5	9.7	1531	2 T42218	slit-1 protein hom
9	360	9.6	907	2 JG0176	orphan G protein-c
10	346	9.3	1523	2 T13953	MEG5 protein - ra
11	344	9.2	603	2 JCT1282	insulin-like growt
12	342.5	9.2	603	2 JCT128	insulin-like growt
13	324	8.7	1469	2 B36655	insulin-like growt
14	324	8.7	1480	2 A36655	slit protein 2 pre
15	321	8.6	560	2 A60164	platelet membrane
16	317.5	8.5	702	2 T21148	hypothetical prote
17	302	8.1	1025	2 T42626	secreted leucine-r
18	299.5	8.0	361	2 A53860	chondroaderin pre
19	298	8.0	536	2 A54901	lysine carboxypept
20	298	8.0	680	2 T19939	hypothetical prote
21	295	7.9	738	2 T19938	hypothetical prote
22	290.5	7.8	1119	2 A01822	leucine-rich repea
23	289.5	7.8	594	2 T23841	hypothetical prote
24	287.5	7.7	458	2 T19941	hypothetical prote
25	286.5	7.7	961	2 T23395	hypothetical prote
26	285.5	7.7	610	2 T23836	hypothetical prote
27	285.5	7.7	1385	2 T13887	tlr protein - trul
28	283	7.6	1066	2 T15864	hypothetical prote
29	283	7.6	1389	2 T13852	gene wheeler prote

30	280.5	7.5	662	2 S42799	garp precursor - h
31	276.5	7.4	357	2 S24317	decorin precursor
32	274.5	7.4	575	2 T29972	hypothetical prote
33	273	7.3	1112	2 T10504	disease resistance
34	272	7.3	1134	1 A29944	oncofetal tropobol
35	270.5	7.2	420	2 A53531	decorin precursor
36	268	7.2	354	2 A55454	decorin precursor
37	268	7.2	360	2 S06280	decorin precursor
38	266	7.1	1097	2 A29943	toll protein precu
39	266	7.1	1535	2 S46224	peroxidasin - trul
40	263	7.0	1016	2 T30553	disease resistance
41	262.5	7.0	562	2 T34319	hypothetical prote
42	262.5	7.0	1051	2 T13174	gp150 protein - fr
43	261	7.0	359	1 NBHUC8	decorin precursor
44	261	7.0	360	2 I47020	decorin - rabb3t
45	258	6.9	354	2 S29145	decorin precursor

#### ALIGNMENTS

RESULT 1  
JCT763  
neuronal leucine-rich repeat protein-3 - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C/Accession: JCT763  
R/Fukamachi, K.; Matsuo, Y.; Kitahara, C.; Kuchino, Y.; Tsuda, H.  
Biochem. Biophys. Res. Commun. 287, 257-263, 2001  
A/Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the g  
A/Reference number: JCT763; PMID:11549284  
A/Contents: Fibrosarcoma cells  
A/Accession: JCT763  
A/Molecule type: mRNA  
A/Residues: 1-707 <FUK>  
A/Cross-references: GB:AF291437  
A/Comment: This protein, a new member of the neuronal leucine-rich repeat protein fa  
in protein-protein interaction and functions as a cell adhesion molecule or soluble  
C/Genetics:  
A/Genes: nlr-3  
C/Keywords: cell adhesion

Query Match	44.2%	Score 1649.5;	DB 2;	Length 707;
Best Local Similarity	47.5%	Pred. No. 4.5e-105;		
Matches 337;	Conservative 112;	Mismatches 229;	Indels 31;	Gaps 7;
QY	6	APLLAMVAG-ATATVPVVPWHPVPCPOCACQIRPYPTRSSYFEATVDCNDLFTAV	63	
DB	4	APLIHVLGLAITALVQAGDKVDCPQCTCEIRPWFTRSYMEASTVDCNDLGLNF	63	
QY	64	PRPLPACTOTLLQSSSYAVVDSSEGLYLANLELDLSONSESDARDCPHALPOLSLH	123	
DB	64	PARLPADTQITLLQTNNTAIIEHS-TDFPVNLTGLDLSQNNLSSVNNINQKMSQLTSY	122	
QY	124	LEENOTLRLEDHSFAGLASQELLYLNHNOXYRTPAFSGSLMLRLHNSNLRRAIDSR	183	
DB	123	LEENKLTLEPEKLYGISNQLLYNNHNLISATSPAFVGLNHLRLHLSNRLQMTNSK	182	
QY	184	WFMLEPNEILTMIGKNKVDAILDMNFRPLANLRSYLAVAGNLRISDYALEGQSSELS	243	
DB	183	WFEALNLEILTMIGDPIRLIKDMNQPLKLSLYIAGINTLEVEDDALVGLNENESIS	242	
QY	244	FYDNOIARVPRALQVPGIKFTLDLKNKPLQRYGPGDFANMLKELGLNNMELYSIDK	303	
DB	243	FYDNRINKVQVALQKAVNLKFLDKKNPDIRTRDGFNMHLKELGINNMEYSIDS	302	
QY	304	FLVYNPELTIKDITNPNPLSFTHPRAFHHLQMETLMMNNNAISLHQQTVSLPNLQE	363	
DB	303	LAVDNLPRDKLTETATNNPRLSYTHPAFFRLKLESLMNSNALSLVHGTTESLNLRK	362	
QY	364	VGLHGPNICDCQVIRVANATGTRVRIEPOSTLCAPPDQLRPLVREVPFREMTHCLPL	423	

```

RESULT      4
JC5239
Insulin-like growth factor acid-labile chain - baboon
C:Species: Papio sp. (baboon)
C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C:Accession: JC5239
R:Delhanty, P.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like
A:Reference number: JC5239; MUID:97040714
A:Contents: liver
A:Accession: JC5239
A:Molecule type: mRNA
A:Residues: 1-605 <DEL>
C:Comment: This factor is structurally related to prolinsulin and have insuline-like meta

```

A:Gene: CESP:121212.9a  
A:Map position: 4  
A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2





R:Botclair, Y.R.; Seto, D.; Hsieh, S.; Hurst, K.R.; Ooi, G.T.  
Proc Natl. Acad. Sci. U.S.A. 93, 10028-10033, 1996  
A:Title: Organization and Chromosomal Localization of the gene encoding the mouse acid 1  
A:Reference number: J66128; MUID:96413591  
A:Accession: J66128  
A:Molecule type: DNA  
A:Residues: 1-603 <B01>  
A:Cross-references: GB:U66900; NID:91621612; PID:AB17270.1; PID:91621613  
C:Comment: This protein is a serum protein and it is of the ternary complex in the phys  
A:Gene: als  
A:Map position: 17

Query Match 9.2%; Score 342.5; DB 2; Length 603;  
Best Local Similarity 22.0%; Pred. No. 1.3e-15;  
Matches 139; Conservative 69; Mismatches 190; Indels 233; Gaps 15;

QY 8 LLLAWA-----GATATVAVVPMHVPCCPOCAQIRWYPRSSRYEATTVDCN 56  
DB 13 LIAFWALGPCYLQGTDPASADAE---GPOCPVCTCSY-----DDYDELSEVCS 61  
QY 57 DEFLAVPALPAGTQTLTLLQNSIVR-----DQSELG 91  
DB 62 SRNLTOLPGICPVSTRALMDGNSSIPSAAFNQISLDELNLQSGSWLRSLPEQALIG- 120  
QY 92 LANTELDSQNSFSDARCDPFHALPOLSLHLENQTLRE----- 133  
DB 121 LQNLVHLERNLRLSLAAGLFPHHPFSLASLSIGNNLLRLLEGFLRGSLHMDLNGN 180  
QY 134 -----DHFAGLASIQEYLNNQLYRAPRFSGLSMLRLHLSNLLRAIDSWFEM 187  
DB 181 SLVLPDVTVOGIGNHLELYLGNKLTLYLQPALCGELREHLSRNALRSVKANVFTH 240  
QY 188 LFNLEILMIGKKNVDI-----LDNN----- 208  
DB 241 LPRLOKLYIRNLITRAVARAFLGKALKRDLDSHNRAVGLLEDFPGLLGHLVRLAHN 300  
QY 209 -----FRPLANLRLSVLAGMNLREISDYALEGSLSESTSFYDNLARV----- 252  
DB 301 AITSLRPRFKDLHLELQGLHNHNRIRQIGKTFEGLQLEVLITLNDNIHEYKVAFG 360  
QY 253 -----PRALEQVPGKFLDLNKNPLQVGPGEFANMLHKEGLINN 295  
DB 361 LFNVAVMNLGNCRLSLRPHVFOGLRHLSTLHESCLGRIRLHPFAGLSGLRLFLRD- 419  
QY 296 BELVSDKFAVNLBELTLDITNNPRLSFIRPRAFHLPQMETLMLNNALSALHQ--- 352  
DB 420 NSISIEROSLAGISELLELDTAN-QLTHLPRLDFOGIGOLEYLLSNNOQLTMSEDVL 478  
QY 353 -----QTVESLPNLOEYGLHGNPIR 372  
DB 479 GPLQAFNIDLSHNLETPAEGLSSGLRKLRLNNSLQTFVPOPGERMLDANPMD 538  
QY 373 CDCVIR-----WANATGRVREIEP-----OSTLCAEPDQLRLPVEVPR 414  
DB 539 CSCPKALRDFALQMPGVPRFVOTVCEGDDCQPYTYTNITCAGPANSGLDLNDI--- 595  
QY 415 EMTDCLPLISRPSPPLQVAGSGSMVLHC 445  
DB 596 -----SETLFVHC 603

RESULT 13  
B3665  
slit protein 2 precursor - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 30-Apr-1991 #sequence: revision 30-Apr-1991 #text: change 19-May-2000  
C:Accession: B3665  
R:Robberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.  
A:Title: Slit: an extracellular protein necessary for development of midline glia and co  
A:Reference number: A3665; MUID:91099665

A:Accession: B3665  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1469 <R07>  
A:Cross-references: GB:X53959  
C:Genetics:

A:Gene: FlyBase:sl  
A:Cross-references: FlyBase:FBgn0003425  
C:Superfamily: unassigned EGF-related proteins; EGF homology; leucine-rich alpha-2-gf

F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>  
F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>  
F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>  
F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>  
F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>  
F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>  
F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>  
F:1028-1061/Domain: EGF homology <EGF2>  
F:1068-1099/Domain: EGF homology <EGF2>  
F:1115-1148/Domain: EGF homology <EGF1>

Query Match 8.7%; Score 324; DB 2; Length 1469;  
Best Local Similarity 24.7%; Pred. No. 9e-14;  
Matches 103; Conservative 67; Mismatches 169; Indels 78; Gaps 11;

QY 29 CPPCACQIRPWYTPRSSRYEATVDCNDLFLFAVPALPAGQTLTLLQNSIVRQDSE 88  
DB 519 CPAMCHC-----EGTYVDCGRKURKEIPRODLPHTLLELDNDELGRISDGG 565  
QY 89 LQYLANTELDSQNSFSDARCDPFHALPOLSLHLENQTLREHDSFAGLSAQEYL 148  
DB 566 L-----FGRPLHVLKLELRQDLGIEPNNAEGASHIDQL 602  
QY 149 NNQLYRIAPRAFGSLNLRHLNSNLRRAIDSWFEMLPNLEILMIGKKNVDAILDN 208  
DB 603 GENKIKESNKKMFLGHQKTLNLYDNOISCVPSFEHLNLSLIMLASNPENC--NCH 660  
QY 209 FRPLAN-LRSLVLAGMNR-----EISDVALGQSLSE----- 240  
DB 661 LAMFECYAKKSLNGAAGCAPSKVYRDQIKDLPHSEKCSSENSEGLGDCYCPSPCT 720  
QY 241 -----SLSEFYDNLARVPRRALEQVPG-LKFLDLNKNPLQVGPGEFANMLHKEGLINN 295  
DB 721 CIGTVAAASRNQLEIPR-----GIPAFSELYLENEHLEQHYERIRHRLSLTRLDLSN- 775  
QY 296 BELVSDKFAVNLBELTLDITNNPRLSFIRPRAFHLPQMETLMLNNALSALHQQT 355  
DB 776 NQITLSVYTFANMLTKLSTLIISYN-KIQCLQRLHALSGLNLRVYSLHGNRISMPEGSF 834  
QY 356 ESLNPLEQVGLHGNPIRDCVIRMANGTGRTREPOSTLCAEPDQLRLPVEVPR 412  
DB 835 EDLKLTHIALGSLNPLVDCGLKMF-SDWIKLDVPEGIARCAEPQKDKLIIISTP 890

RESULT 14  
A3665

C:Comment: The amino end of the intact protein is blocked.  
 C:Comment: This protein is absent in Bernard-Soulier syndrome.  
 C:Genetics:  
 A:Gene: GDB:GPI5  
 A:Cross-references: GDB:230236; OMIM:173511  
 A:Map position: 5pter-5qter  
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology  
 C:Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane prot

Query Match 8.6%; Score 321; DB 2; Length 560;  
 Best local similarity 26.6%; Pred No. 3.5e-14;  
 Matches 141; Conservative 57; Mismatches 227; Indels 106; Gaps 16;

```

QY 9 LIAWAGATATVYVPMWHPQPCACQIRPWYTPRSSYREATTVDCNDLF-LTAVPPAL 67
DB 6 LICAVIGLIRAGP-----PPCPACKC-----VFRDAACSGGDVARISAL--GL 48
QY 68 PAG-TQTLLQSNSTYRVQSELGTYLANTELDLSONSFSDARCDPFAHPQLLSHLEE 126
DB 49 PNLTHILLFGMGQVLOSQSFSG-MTVLQRLMISDSHISAVAGTFSDLIKIKTLRLSR 107
QY 127 NOTRLDEHSPAGIASIOELYNHNLRYRAPAFSGSLNLRHLNSNLRRAIDSRMFE 186
DB 108 NKITHLPGLLDKMYLLEQFLDHNALRGIDQMFQKLVNQLDELALNQDLPLPASLFT 167
QY 187 MLPNTLEILMIGNKVDAIIDMNERPLANRSLVLAGMNLREISDYALEGHSLESFYD 246
DB 168 NLENKLLDLSGNKLTFLPKGLGAQAKLERLLHSNRLVSLDSGLINSIGALTELQFHR 227
QY 247 NOLAVPRRALEQVPGKFLDLNK-----NPLQRVGPGDFA 282
DB 228 NHIRSIAPGAFDRPLPMLSSITLSRNHLAFPSALFLHSNLTLLTFENPLAELPGVLF 287
QY 283 NMLHKEIGLNNEELVSDKFAIVNLPETKIDITNNRSLFIHPRAFHLPQMETML 342
DB 288 EMGGLQELMLNR-TQLRTLPAAFRNLSRLRYLGVTLSPRLSALPGAFQGLGELQVLAL 346
QY 343 NNNALSAL-----HOQ-----T 354
DB 347 HNGSLTALPDGLRGIGLRLROYSLRRNRRLRALPRALFRNLSLESYQLDHNQLETLP 406
QY 355 VESLPMLQEVGLHGNPIKDCD---VIRMANATGTRVREIEPOSTLCAEPPDLQRLPYRE 410
DB 407 FGALPRLTEVLIGHNSWRCDGIGPFLGWLRLQHLGLVGGEPPR--CAGRGAHAGLPMA 464
QY 411 VFPREMTDHCPLDISRSPPSLQVASESMVHCRALAPPEPEIT-WVTP 460
DB 465 LPEGDA--EC---PGRPPPPRPADSSSEAPVH--PALAPNSSEPPVMAOP 509
  
```

Search completed: August 31, 2002, 14:36:40  
 Job time: 102 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:35:28 ; Search time 13.6 Seconds

(without alignments)  
2029.929 Million cell updates/sec

Title: US-09-905-088a-245

Sequence: 1 MRLVAPLLLAIVAGATATV.....RKLPSSGERTLPLPLSGNS 713

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3718	99.6	713	1	GAC1_HUMAN
2	407.5	10.9	605	1	AUS_HUMAN
3	397.5	10.7	605	1	ALS_PAPHA
4	344	9.2	603	1	ALS_RAT
5	342.5	9.2	603	1	ALS_MOUSE
6	340.5	9.1	567	1	GPV_MOUSE
7	324	8.7	1480	1	SLIT_DROME
8	321	8.6	560	1	GPV_HUMAN
9	314	8.4	905	1	TLR3_MOUSE
10	313.5	8.4	567	1	GPV_MOUSE
11	298.5	8.0	361	1	CHAD_BOVIN
12	298	8.0	536	1	CAPE8_HUMAN
13	287	7.7	904	1	TLR3_HUMAN
14	280.5	7.5	662	1	GNP_HUMAN
15	279	7.5	782	1	CHAO_TRICA
16	276.5	7.4	357	1	PGS2_CHICK
17	272	7.3	1315	1	CHAO_DROME
18	270	7.2	360	1	PGS2_PIG
19	270	7.2	1050	1	TLR7_MOUSE
20	269.5	7.2	1032	1	TLR9_MOUSE
21	268.5	7.2	1032	1	TLR9_HUMAN
22	268	7.2	354	1	PGS2_MOUSE
23	268	7.2	354	1	PGS2_BOVIN
24	266	7.1	1097	1	TOLL_DROME
25	261	7.0	359	1	PGS2_HUMAN
26	261	7.0	360	1	PGS2_RABIT
27	260	7.0	977	1	Y848_HUMAN
28	259	6.9	331	1	PLIB_AGBXL
29	258	6.9	354	1	PGS2_RAT
30	253.5	6.8	1049	1	TLR7_HUMAN
31	253	6.8	360	1	PGS2_CANFA
32	253	6.8	372	1	PGS1_HORSE
33	253	6.8	1032	1	TLR8_MOUSE

34	251.5	6.7	369	1	PGS1_SHEEP	046390 ovis aries
35	251	6.7	360	1	PGS2_HORSE	046542 equus caball
36	250.5	6.7	369	1	PGS1_BOVIN	P21809 bos taurus
37	250.5	6.7	682	1	CONN_DROME	001819 drosophila
38	249	6.7	369	1	PGS1_MOUSE	P28653 mus musculus
39	248.5	6.7	369	1	PGS1_CANFA	002678 canis famli
40	248	6.6	369	1	PGS1_RAT	P47853 rattus norv
41	246.5	6.6	368	1	PGS1_HUMAN	P21810 homo sapien
42	246.5	6.6	2026	1	CYAA_YEAST	P08678 saccharomyc
43	242.5	6.5	582	1	SHO2_HUMAN	Q9ug13 homo sapien
44	242.5	6.5	582	1	SHO2_MOUSE	088520 mus musculu
45	241	6.5	342	1	LDM_BOVIN	Q05443 bos taurus

## ALIGNMENTS

RESULT	ID	STANDARD	PRT	713 AA
1	GAC1_HUMAN			
AC	075325;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Glioma amplified on chromosome 1 protein precursor.			
GN	GAC1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Glial tumor;			
RX	MEDLINE=98324709; PubMed=9662332;			
RA	Malloy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Mulleris M.,			
RA	Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;			
RT	"GAC1, a new member of the leucine-rich repeat superfamily on			
RT	chromosome band 1q32.1, is amplified and overexpressed in malignant			
RT	gliomas."			
RL	Oncogene 16:2997-3002(1998).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).			
CC	-1- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.			
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: AF030435; AAC39792.1; -			
DR	MIM: 605492; -			
DR	InterPro: IPR003006; IG_MHC.			
DR	InterPro: IPR003598; IG_C2.			
DR	InterPro: IPR001611; LRR.			
DR	InterPro: IPR000483; LRR_Nterm.			
DR	InterPro: IPR000372; LRR_Cterm.			
DR	InterPro: IPR003592; LRR_out.			
DR	InterPro: IPR003591; LRR_Typ.			
DR	Pfam: PF00047; Ig_1.			
DR	Pfam: PF00560; LRR; 10.			
DR	Pfam: PF01463; LRRCT; 1.			
DR	Pfam: PF01462; LRRNT; 1.			
DR	PRINTS: PR00019; LRRICHPT.			
DR	SMART: SM00408; IGC2; 1.			
DR	SMART: SM00370; LRR; 6.			
DR	SMART: SM00082; LRRCT; 1.			
DR	SMART: SM00013; LRRNT; 1.			
DR	SMART: SM00369; LRR_Typ; 2.			

```

DR SMART; SM000369; LRRNT; 1.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605
FT REPEAT 53 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 145 168
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 313 336
FT REPEAT 337 360
FT REPEAT 361 384
FT REPEAT 386 408
FT REPEAT 409 432
FT REPEAT 433 456
FT REPEAT 458 480
FT REPEAT 482 504
FT REPEAT 505 530
FT CARBOHYD 64 64
FT CARBOHYD 85 85
FT CARBOHYD 96 96
FT CARBOHYD 368 368
FT CARBOHYD 515 515
FT CARBOHYD 580 580
SO SEQUENCE 605 AA; 66034 MW; F6562A23CBE918F6 CRC64;

Query Match 10.9%; Score 407.5; DB 1; Length 605;
Best Local Similarity 25.2%; Pred. No. 6; Se-22;
Matches 152; Conservative 72; Mismatches 173; Indels 205; Gaps 16;

QY 3 LTVAPLLAWVA-----GATATVYVWVWHPCCPOCACQIRPWTPRSSY---REATT 52
DB 8 LALALLLSWVALGPRSLGADPGTGEAGPACRACAC-----STDDADELS 57
QY 53 VDCNLEFLTAVPALPAGTQITLLSNSIVRY-----DQSELGY----- 91
DB 58 VFCSSRNLTFLPDGVPGTQALMDGNMNSVPPAPAFQMLSSGFNLGQGLSLEPGA 117
QY 92 ---LANLTFLDSONSFSDARCDPFHALPOLSLHEENQTLRL----- 133
DB 118 LLGLENLCHLHERNOLNSLAGTFAHTPAPALASLGISNRLSLRLEGLGEGLSIMDLNL 177
QY 134 -----DHSFAGIASLOELTYLNHNOLYRIAPRAFSGLSMLRLHLNSLRAIDSRW 184
DB 178 GWNSLAVLPDAAFRGISRELTYLAGNRLAYQPALFSGSLARELDSRNALRAIKAVY 237
QY 185 FEMAPNLEILMIGGKKVDAI-----LDMN----- 208
DB 238 FVQLPRLQKLYLDRMLIAVAGAFGLKALRWLDLSNHRVAGLLEDPFGLLGLRYLRL 297
QY 209 -----FRPLANLRSLVLAGMNLRESDVALSGLSLESLSTYDNOIARV---- 252
DB 298 SHNAIASLRPTTFKDLHLLEELQLSGHNRIROLAERSFEGQLQEVLTLDHNOIQEVKGA 357
QY 253 -----PRRALEQVPGKFLDLNKNPRLQVRGPGDFANMLHLKELGL 292
DB 358 FLGLTNVAVMNLGNCRLNLPBQVFRGKGLSHLSLBSSCLRIAPHTFTGSLRRLFL 417
QY 293 NMMEELVSIKFAVLNDELTKIDITNNPRLSFTHPRAFHHLPOMETIMLNNALSAI-- 350
DB 418 KD-NGIVGIEBQSLMGALAELELDITSN-QTLHPRLHRLFOGKGLKLYLLSSNRLAEIPA 475
QY 351 -----HOQTVESLPN-----LOEVGLHG 368

```

```

DB 476 DALGPIQRAFWLDVSHNRLEALPNLSLLAPLGRRLYLISLRNNSLRTPPPGGLERLMEG 535
QY 369 NPTRCCVIR-----WANAQTRY-RFIEP-----QSTICAPPDQLRPLVR 409
DB 536 NPWDCCGPELALRDFALQNSAVPRFVQALCEGDGCPAPAYTYNITCASPPVEVGLDR 595
QY 410 EV 411
DB 596 DL 597

RESULT 3
ALS_PAPHA
ID ALS_PAPHA STANDARD; PRT; 605 AA.
AC 002633;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
XP [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=97040714; PubMed=8886027;
RA Delhanty P., Baxter R.C.;
RT "The cloning and expression of the baboon acid-labile subunit of the
RT Insulin-like growth factor binding protein complex.";
RL Biochem. Biophys. Res. Commun. 227:897-902(1996).
CC -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
CC IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
CC IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S83462; -; NOT_ANNOTATED_CDS.
DR HSSP; P23945; LXUN.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 19.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEORICHRPT.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRRNT; 1.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605
FT REPEAT 53 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 168 168

```

```

FT REPEAT 194 216 LRR 7.
FT REPEAT 217 240 LRR 8.
FT REPEAT 242 264 LRR 9.
FT REPEAT 266 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 337 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 478 LRR 18.
FT REPEAT 479 504 LRR 19.
FT REPEAT 506 529 LRR 20.
FT REPEAT 543 566 LRR 21.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64;

```

Query Match 9.2%; Score 344; DB 1; Length 603;  
 Best Local Similarity 23.3%; Pred. No. 2.6e-17;  
 Matches 141; Conservative 63; Mismatches 194; Indels 208; Gaps 16;

```

OY 8 LILAWA-----GATATPVVPMVHVPPOCACQIRWYPRSSYREATTVDN 56
   |||||
DB 13 LIAFWALGBCHLQGTDPGASDAE---GPQCPVACTCS-----HDDYTDLSVFC 61
   |||||
OY 57 DFLTAVPAPLPAQGTOTLLQNSIVRV-----DQSLG 91
   |||||
DB 62 SKNLTHLPDPIPVSTRALMDGNMSSIPSAQONSLDFLWQSGWLSRLEPQALG- 120
   |||||
OY 92 LANLTELDSQNSFSDARDCDFALPOLLSLHEENOLRLE----- 133
   |||||
DB 121 LQNLTYLHERNRLRLAVGLFTHPTSLASLSSNLLGLEGLPOGSLHMDLNGWN 180
   |||||
OY 134 -----DHSEFAGIASQELLYLNHNOYLRIAPRAFSGLSNLRHLHNSNLRADSRFEM 187
   |||||
DB 181 SLVVLPTDVFQGLGNLHLYLAGNKLTYLQPALFCGLGELRELDLSRNALRSYKAVFVH 240
   |||||
OY 188 LFNLELIMIGKNVDAT-----LDNN----- 208
   |||||
DB 241 LPRQLKLYIDRLNLTAVAGAPLGMALMKLDSLHNRVAGLMEDTPGLGLHVLRLAHN 300
   |||||
OY 209 -----FRPLANLRSVLVLAGMNLREISDYALBGLSQLESLSFYDNOIARV----- 252
   |||||
DB 301 AIASLRPRFTKDLHLEELQGNHNRIRQGERFEGLQLEVLITLNDQITEVRVAFSG 360
   |||||
OY 253 -----PRALAEQVGLKFLDLINKNPLORVGGDFANMLHKEGLGNM 295
   |||||
DB 361 LFNVAVMNLSGNCLSLPRVVFQGLDKLHSLHESCLGHVRLHPTFAGLSGLRFLRD- 419
   |||||
OY 296 BELVSIKFAVLNLPFLTLDITNNPRLSFIHPRAFHHLPOWETLMLNNALSAL----- 350
   |||||
DB 420 NSISSIEROSLAGLSLELDTLN-RLTHLPRLQFQGLHLEYLLSYNOITLTSAEVL 478
   |||||
OY 351 -----HQ-----QVESLPNQEVGLHGNPRL 372
   |||||
DB 479 GLQRAFMLDISHNLETLAEGLSLGNVRLSLRNNLSQFFSPQGLERLMLDANPMD 538
   |||||
OY 373 CDCVLR-----WANNATGTRVRIEP-----OSTLCAEPDRLQRLPVREVPRL 414
   |||||
DB 539 CSCPLKALRDFALQNPVGVPRVQTVCEGDDCPVYTYNNITCAGPANVSGLDLRDVSST 598
   |||||
OY 415 EMTDHC 420
   ||
DB 599 HFV-HC 603

```

```

RESULT 5
ALS_MOUSE
ID ALS_MOUSE STANDARD; PRT; 603 AA.
AC P70389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS OR ALBS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=129/SV;
RX MEDLINE=96413591; PubMed=8816745;
RA Boiscclair Y.R., Seto D., Hsieh S., Hurst K.R., Ooi G.T.;
RT "Organization and chromosomal localization of the gene encoding the
RT mouse acid labile subunit of the insulin-like growth factor binding
RT complex."
RL Proc. Natl. Acad. Sci. U.S.A. 93:10028-10033(1996).
CC -1- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
CC CIRCULATING IGFs TO THE TISSUES.
CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
CC IGF-1 OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U66900; AAB17270.1; -.
DR MGD; MGI:107973; Igfals.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR; 19.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 9.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 603
FT REPEAT 52 73
FT REPEAT 74 96
FT REPEAT 98 120
FT REPEAT 121 144
FT REPEAT 146 168
FT REPEAT 169 192
FT REPEAT 194 216
FT REPEAT 217 240
FT REPEAT 242 264
FT REPEAT 265 288
FT REPEAT 289 312
FT REPEAT 312 336
FT REPEAT 338 360
FT REPEAT 361 384
FT REPEAT 386 408
FT LRR 1.
FT LRR 2.
FT LRR 3.
FT LRR 4.
FT LRR 5.
FT LRR 6.
FT LRR 7.
FT LRR 8.
FT LRR 9.
FT LRR 10.
FT LRR 11.
FT LRR 12.
FT LRR 13.
FT LRR 14.
FT LRR 15.

```

SO SEQUENCE 567 AA; 63467 MW; C48643AA73967A7D CRC64;

Query Match	9.1%;	Score 340.5;	DB 1;	Length 567;
Best Local Similarity	24.6%;	Pred. NO. 4.2e-17;		
Matches 132;	Conservative 70;	Mismatches 217;	Indels 117;	Gaps 15;

```

0Y 17 TATVAVV---PMHNDPCPOCAOQR-----PWYT 42
Db 8 SAVLEPLAAOPR--PCPPTCKCVVDAQAQSGSGVAHIAELGPTNLTHILFERMOGIL 65
0Y 43 PRSSYREATYVD---CNDFLTAVPA-----LPAG-----TQ 72
Db 66 RNHSSEGGTYVQROMLSHSHSALIDPGFENDLYVKTLRLTRNKISRLPRAILDKWYLE 125
0Y 73 TLLQNSIVRVDQSEGLYLANITELDI,SONSFSDARCDGFHALPOLLSHLEENOTRL 132
Db 126 QLFIDHNALROLDQNLFOQLNNIOELGNOQNSFLRANFESSIREKLIDLSRNNITLH 185
0Y 133 EDHSAPAGIASIQEILYNHNOYLRIAPRAFSGLSNULRLHINSULRAIDSREMTLPNE 192
Db 186 PKGLIGAQVLEKLEKILLYSNQLTSDVSGHLSNIGALTLEIRLERNHRSVARGADRLGNLS 245
0Y 193 ILMIGNNVDAI,IDMNRPLANLSYVLACNNLREISDYALEGQSIES,SPFDNOARY 252
Db 246 SLTISGNLLESLPRALPLAHVSVAIRLTLFENRPELIDYVLEGEMAGIRELMNGTHLSTL 305
0Y 253 PRRALEOYPGLEKFLDKNKP--LQVWGGDFANMHLRELGINNMEEVSDIKFALVNLPE 311
Db 306 PRAAFRNISGIGTGLTRNPRSLALPRGVOGLELRYVLDLH--NALABELRDDALGELG 364
0Y 312 LTKLIDTINPRLSFIHPRAFHNHLPOMETLM,NNNALASAL,NOQTVESL,PNLQOEVGLHGPNP 371
Db 365 LROYSLRHN--RLRALPRTLFRNLISLESBVOLEHNQETLETDGVFALPQLTQVLLGHNP 423
0Y 372 RQDC-----VIFW-----ANAGTNYKRTLEPOST-----CAEPDLORLPV 408
Db 424 LQDGLMFLQWLWNRHHPDILGRDEPPQCRBEPRASLSJFWELLQDGPWCPDP----- 475
0Y 409 REVFREKTDLCPLRISPSFPPLQ-----VAGSEVNLGRALAEPEPIYW 457
Db 476 RSLPDLPEPTENALAPYPSMLPNSWQOTYANQVARGES-----PNNRLYW 521

```

```

RESULT 7
SLIT_DROME
ID SLIT_DROME STANDARD; PRT; 1480 AA.
AC P24014;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Slit protein precursor.
GN SLIT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91099665; Pubmed=2176636;
RA Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
RT "Slit, an extracellular protein necessary for development of midline
RT glia and commissural axon pathways contains both EGF and LRR
RT domains.";
RL Genes Dev. 4: 2169-2187(1990).
CC -1- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
CC MATRIX MOLECULES.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.

```

```

CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53959; CAA37910.1; -.
DR PIR; A36665; A36665.
DR HSSP; P00743; ICGF.
DR FlyBase; FBgn0003425; sl1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR000561; EGF_1like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00008; EGF; 7.
DR Pfam; PF00054; Laminin_G; 1.
DR Pfam; PF00560; LRR; 17.
DR Pfam; PF01463; LRRT; 4.
DR Pfam; PF01462; LRNT; 4.
DR SMART; SM00041; CT; 1.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF_1like; 5.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRT; 4.
DR SMART; SM00013; LRNT; 4.
DR SMART; SM00369; LRR_Typ; 9.
DR SMART; SM00282; Lamg; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PSS0025; LAM_G_DOMAIN; 1.
KW Neurogenesis; Glycoprotein; Signal; Alternative splicing;
KW EGF-like domain; Repeat; Leucine-rich repeat.
FT SIGNAL 1 36
FT CHAIN 37 1480 SLIT PROTEIN.
FT REPEAT 99 122 LRR 1.
FT REPEAT 123 146 LRR 2.
FT REPEAT 148 170 LRR 3.
FT REPEAT 171 194 LRR 4.
FT REPEAT 195 218 LRR 5.
FT REPEAT 220 246 LRR 6.
FT REPEAT 321 344 LRR 7.
FT REPEAT 345 368 LRR 8.
FT REPEAT 369 392 LRR 9.
FT REPEAT 394 416 LRR 10.
FT REPEAT 417 440 LRR 11.
FT REPEAT 522 545 LRR 12.
FT REPEAT 546 569 LRR 13.
FT REPEAT 570 593 LRR 14.
FT REPEAT 595 617 LRR 15.
FT REPEAT 618 641 LRR 16.
FT REPEAT 643 666 LRR 17.
FT REPEAT 678 701 LRR 18.
FT REPEAT 720 743 LRR 19.
FT REPEAT 745 764 LRR 20.

```

CC	-1	TISSUE SPECIFICITY: PLATELETS AND MEGAKARYOCYTES.	
CC	-1	PM: THE N-TERMINUS IS BLOCKED.	
CC	-1	SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).	
CC		-----	
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC		the European Bioinformatics Institute. There are no restrictions on its	
CC		use by non-profit institutions as long as its content is in no way	
CC		modified and this statement is not removed. Usage by and for commercial	
CC		entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).	
CC		-----	
DR	EMBL: L11238; AAA03069.1; -;		
DR	EMBL: Z23091; CAA80637.1; -;		
DR	HSSP; P16473; 1XOM.		
DR	MIM; 173511; -;		
DR	InterPro: IPR001611; LRR.		
DR	InterPro: IPR000483; LRR_Cterm.		
DR	InterPro: IPR000372; LRR_Nterm.		
DR	InterPro: IPR003592; LRR_Out.		
DR	InterPro: IPR003591; LRR_Typ.		
DR	Pfam; PF00560; LRR; 14.		
DR	Pfam; PF01463; LRRCP; 1.		
DR	PRINTS; PRO0019; LEUCRICRPT.		
DR	SMART; SM00370; LRR; 1.		
DR	SMART; SM00082; LRRCT; 1.		
DR	SMART; SM00369; LRR_Typ; 9.		
KW	Platelet; Transmembrane; Glycoprotein; Blood coagulation;		
KW	Repeat; Leucine-rich repeat; Cell adhesion; Signal.		
FT	SIGNAL	1	16
FT	CHAIN	17	560
FT	DOMAIN	17	523
FT	TRANSERM	524	544
FT	DOMAIN	545	560
FT	REPEAT	73	96
FT	REPEAT	97	120
FT	REPEAT	122	144
FT	REPEAT	145	168
FT	REPEAT	169	192
FT	REPEAT	194	216
FT	REPEAT	217	240
FT	REPEAT	241	264
FT	REPEAT	266	288
FT	REPEAT	289	312
FT	REPEAT	314	337
FT	REPEAT	338	361
FT	REPEAT	362	385
FT	REPEAT	386	409
FT	REPEAT	409	509
FT	CARBOHYD	51	51
FT	CARBOHYD	181	181
FT	CARBOHYD	243	243
FT	CARBOHYD	267	267
FT	CARBOHYD	298	298
FT	CARBOHYD	312	312
FT	CARBOHYD	385	385
FT	CARBOHYD	499	499
FT	CONFLICT	73	74
FT	CONFLICT	109	109
FT	CONFLICT	130	130
FT	CONFLICT	136	138
FT	CONFLICT	209	209
FT	CONFLICT	267	267
FT	CONFLICT	327	327
FT	CONFLICT	478	478
FT	CONFLICT	509	509
FT	SEQUENCE	560 AA; 60959 MW; 60959 MM; 60959 MC64;	

QY	9	LLAMAGATATVPVPMVHPPPOCACQIRWYTPRSTRGATVDCNDLF-LTAAVPPAL	67
QY	9	LLAMAGATATVPVPMVHPPPOCACQIRWYTPRSTRGATVDCNDLF-LTAAVPPAL	67
Db	6	LLCAVLGLLRAP-----FPPPPACK-----VFRAAACSGGDVARIAL--GL	48
QY	68	PAG-TOTLLLOSNIVRVDSELSGLYLANLTETLDSONSPFAROCDFHALPOLSLHLE	126
Db	49	PTNLTHILLPFMGKGVQLQSSFSG-MVYLQKMLISDSHSAVAGTFSDLIKKTLNLSR	107
QY	127	NOLTRLEDHSPAGLASLOELYLNHNOLYRIAPRAFSGLSNLRHLNLSLRAIDSRWF	186
Db	108	NKITHLPALLDKVWLLLEOLFELDHNAALRGIDQNNFQKLVNLQELALNONDELFPASLFT	167
QY	187	MLPNEILMIGGNVDAIDLNNFRPLANLRSVLACGNLRISDYALEGLQSLESFYD	246
Db	168	NLENKTLIDSGNNLTHTLPRKGLGAQAOKLERLLHNSKLVSLDSGLNLSGALTELFHR	227
QY	247	NOLARVPRALEQVPGKLFEDLNK-----NPLQRYGPDFA	282
Db	228	NHISIAIAGADRLPRNLSSLLTSKNHLAFELPSALFLSHNLTLTLENPLAELPGVLF	287
QY	283	NMLHLKELGNMKEELVSIKFAVLNLPETLKIDITNNPLSFTHPPAHHLPOMETIML	342
Db	288	EMGQELMELNLR-TQLRFLPAAPRNLRSRLRYLGVTLSPRLSALPGCAFQGLGELYAL	346
QY	343	NNNNLSAL-----HQ-----T	355
Db	347	HSNGTLAPDGLGLGKLRQVSLRRNRRLRALPRALFRNLSLESVOLDHQLETLPGDV	406
QY	355	VESLPNLOEVLGHNPIKDCD---VIRMANAVTGRVFIEPQSTGLCAEPDQLRPVRE	410
Db	407	FGALPRLETVLLGHNNSRCDGCLPFLGWLROHGLGVGGEPPR--CAGGCAHAGLPLMA	466
QY	411	VPFEMTDHCLPLISPRSFPSTLOVAGSGSNVYLHCALAEPEEITY-WVTP	460
Db	465	LPGGDA--EC---GPPRGPPPPRAADSSSEAPVH-PALAPNSSEPPWMAQ	509
RESULT	9		
TLR3_MOUSE	TLR3_MOUSE	STANDARD;	PRT; 905 AA.
AC	099MB1; 0912M4;		
DT	01-MAR-2002 (Rel. 41, Created)		
DT	01-MAR-2002 (Rel. 41, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Toll-like receptor 3 precursor.		
GN	TLR3.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=129/SV;		
RX	PubMed=11607032;		
RA	Alexopoulos L., Holt A.C., Medzhitov R., Flavell R.A.;		
RT	"Recognition of double-stranded RNA and activation of NF-kappaB by		
RT	Toll-like receptor 3.";		
RL	Nature 413:732-738(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALB/c x NIH; TISSUE=Macrophage;		
RA	Appelquist S.E., Ljunggren H.G.;		
RT	"Molecular cloning of mouse Toll-like receptor 3 cDNA.";		
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: Participates in the innate immune response to microbial		
CC	agents. May be involved in the recognition of ds-RNA. Acts via		
CC	MyD88 and TRAF6, leading to NF-kappa-B activation, cytokine		
CC	secretion and the inflammatory response (By similarity).		
CC	-!- SUBUNIT: Binds MyD88 via their respective TIR domains (By		
CC	similarity).		
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).		
CC	-!- TISSUE SPECIFICITY: Highly expressed in lung. After		
CC	intraperitoneal injection of lipopolysaccharide, highly expressed		



```

CC      INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A
CC      CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Z69596; CAA93440.1; -
DR      InterPro: IPR001611; LRR.
DR      InterPro: IPR000483; LRR_Cterm.
DR      InterPro: IPR000372; LRR_Nterm.
DR      InterPro: IPR003592; LRR_out.
DR      InterPro: IPR003591; LRR_typ.
DR      Pfam: PF00560; LRR_14.
DR      Pfam: PF01463; LRRCT; 1.
DR      PRINTS: PR00019; LEURICHRPT.
DR      SMART: SM00370; LRR; 2.
DR      SMART: SM00082; LRRCT; 1.
DR      SMART: SM00013; LRRNT; 1.
DR      SMART: SM00369; LRR_typ; 10.
KW      Platelet; Transmembrane; Glycoprotein; Blood coagulation;
KW      Repeat; Leucine-rich repeat; Cell adhesion; Signal.
FT      SIGNAL 1 16
FT      CHAIN 17 567
FT      DOMAIN 17 522
FT      TRANSMEM 523 543
FT      DOMAIN 544 567
FT      REPEAT 73 96
FT      REPEAT 97 120
FT      REPEAT 122 144
FT      REPEAT 145 168
FT      REPEAT 169 192
FT      REPEAT 194 216
FT      REPEAT 217 240
FT      REPEAT 241 264
FT      REPEAT 266 288
FT      REPEAT 289 312
FT      REPEAT 314 337
FT      REPEAT 338 361
FT      REPEAT 362 385
FT      REPEAT 387 409
FT      CARBOHYD 51 51
FT      CARBOHYD 181 181
FT      CARBOHYD 243 243
FT      CARBOHYD 298 298
FT      CARBOHYD 312 312
FT      CARBOHYD 385 385
FT      CARBOHYD 498 498
SQ      SEQUENCE 567 AA; 63344 MW; CAl0708ED03707F CRC64;

```

Query Match 8.4%; Score 313.5; DB 1; Length 567;  
 Best Local Similarity 24.0%; Pred. No. 3 8e-15;  
 Matches 133; Conservative 62; Mismatches 226; Indels 133; Gaps 13;

```

QY 4 LVAPLLAVAGATATVPVPHVPCPOACQIRPWYPRSSYREATVDNDLFLTAV 63
DB 1 MRSVLLASVLSVGAQP-----FPCPKTKCYVR-----DAVQSGGSVAHI 43
QY 64 PP-ALPAG-----TQTLLOSNSIVRVDQSELGLIANLTEL 98
DB 44 AELGLPTNLTLLIPMDRGVLSQSHSFGMTVLRMLSDSHSAIDPGTFNDLVKLKTL 103
QY 99 DLSQNSFSDARDQDFHALPOLLSLHLEENQLTFLSDHSFAGLASIDELVLNHNQLRINP 158
DB 104 RLTRNKISHLPAIILDKMVLILQGLFDHNAALRDLDONLFQKLNLKDLCLNONQLSELP 163

```

```

QY 159 RAESGLNLTSLHNSNLLRAIDSRWFEMLPNLEIMIGNKVDAILDMNRPRLSL 218
DB 164 NLFSSGLKVLDSLRNNLTSLPGLGAGQTKLEKLLYSNR---LMSLSGLIANGL 220
QY 219 V---LAQNLNREISDYALEGSLSESIFYDNOLARVRRALEQVPGKFLDKNKLOR 275
DB 221 TELTERNHLSRIAPGAFDSIGNLSTLISGNLLESPPALFLHVSMTLTLTFENLEE 280
QY 276 VGPDPFANMLHKLGLGNMNEELVITKFLVNLPELTKLDITNPLSTIHPRAFHLP 335
DB 281 LPEVLFEEMKGLRELWTNG--THKRTLPAAFRNLISGLQGLTRNPLLSALPGEKFGHLT 339
QY 336 QMETLMLNNALSAHQOFTESLPNLOEVGHGNPVR----- 372
DB 340 ELRLAVHTNLTAELEPDALRGRLRQVSLRNRRLALPRTLFRNLSSLVTVOLEHNOL 399
QY 373 -----CDC-----YIRMANATGTRVRFTEPOSTCAEPDOL 403
DB 400 KTLPGDVFAALPOLTRVLLGHNPMLCDGLMPFLQWLRHHLLELGRDEPQ--CNGPESR 457
QY 404 QRLPV-----REVPRKMDHCLPLISPSFSPSLQ-----VASGESMTL 443
DB 458 ASLTFEMLLOGDQWCPSRRGLPPDPPTENALKAADPTORPNSGQSWAVQLVARGES--- 514
QY 444 HCRALAEPEPEIYW 457
DB 515 -----PDNRFTW 521

```

RESULT 11  
 CHAD\_BOVIN STANDARD; PRT; 361 AA.  
 ID 027972;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Chondroadherin precursor (Cartilage leucine-rich protein) (38 kDa bone  
 DE protein).  
 GN CHAD.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Cartilage.  
 RX MEDLINE=94342341; PubMed=8063792;  
 RA Neame P.J., Sommarin Y., Boynton R.E., Heinegaard D.;  
 RT "The structure of a 38-kDa leucine-rich protein (chondroadherin)  
 RT isolated from bovine cartilage."  
 RL J. Biol. Chem. 269:21547-21554(1994).  
 RN [2]  
 RP SEQUENCE OF 25-55 AND 77-97.  
 RC TISSUE=bone;  
 RX MEDLINE=95113864; PubMed=7814406;  
 RA Hu B., Coulson L., Moyer B., Price P.A.;  
 RT "Isolation and molecular cloning of a novel bone phosphoprotein  
 RT related in sequence to the cystatin family of thiol protease  
 RT inhibitors."  
 RL J. Biol. Chem. 270:431-436(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 008018; AAA21330.1; -  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Cterm.



DB 330 IPDSFLEHLSLCLFNLNRCLTFEARRLGSLPCMLMLDLSSNALETLELGAHALGSLR 389  
QY 276 -----VGPDPFANMLHKL-----GLNMEELY 299  
DB 390 TLLGNALRDLPPYTFYFANLASIQLNLGNRVSPPGGPDPSPGCVAFSGITSLRSL 449  
QY 300 SLDK-----PALVLPETLTKIDITNNP-----321  
DB 450 LVDNEIELLRAGAFHTP-LTELDLSSNPGLEVAATGALGSLASLEVALQGNLWLOY 508  
QY 322 -----RLSIHPRFHHLP-----QMETLMLNNALSLHOOYVSL-PNIOEVG 365  
DB 509 DLPCFICLRNLNAENR-LSHPAWTQAVSLVDLRNNSFILLPGSAMGLETSIRLRY 567  
QY 366 LAGNPICRDCVIRW-ANATGFRFRFIEPOSTIC 397  
DB 568 LQGNPLSC-CGNGWLAOLHOGRVDAVDATODLIC 600

RESULT 15  
CHAO\_TRICA STANDARD; PRT: 782 AA.  
AC P82963; 16-0CT-2001 (Rel. 40, Created)  
DT 16-0CT-2001 (Rel. 40, Last sequence update)  
DT 16-0CT-2001 (Rel. 40, Last annotation update)  
DE Chaoptin (Photoreceptor cell-specific membrane protein) (Fragment).  
GN CHP  
OS Tribolium castaneum (Red flour beetle).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
OC Cucujiforma; Tenebrionidae; Tribolium.  
OX NCBI\_TaxID=7070;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GA-1;  
RA Nie W., Stronach B., Panganiban G., Brown S., Denell R.;  
RT "Molecular characterization of Tolabial and the 3' end of the  
RT tribolium homeotic complex";  
RL submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: REQUIRED FOR PHOTORECEPTOR CELL MORPHOGENESIS. MEDIATES  
CC HOMOPHILIC CELLULAR ADHESION (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR SURFACE OF R-CELL PLASMA  
CC MEMBRANE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CHAOPTIN FAMILY.  
CC -1- SIMILARITY: CONTAINS AT LEAST 21 LEUCINE-RICH REPEATS (LRR).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF32227; AAK01654.1;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR003591; LRR\_Lyp.  
DR Pfam; PF00560; LRR; 19.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 6.  
DR SMART; SM00369; LRR\_Lyp; 14.  
KM Glycoprotein; Membrane; Repeat; Leucine-rich repeat; Vision.  
FT NON\_TR 1  
FT REPEAT 43 66 LRR 1.  
FT REPEAT 67 90 LRR 2.  
FT REPEAT 93 116 LRR 3.  
FT REPEAT 117 140 LRR 4.  
FT REPEAT 141 164 LRR 5.  
FT REPEAT 165 188 LRR 6.  
FT REPEAT 224 247 LRR 7.

FT REPEAT 249 272 LRR 8.  
FT REPEAT 273 296 LRR 9.  
FT REPEAT 297 320 LRR 10.  
FT REPEAT 321 343 LRR 11.  
FT REPEAT 344 367 LRR 12.  
FT REPEAT 370 391 LRR 13.  
FT REPEAT 395 418 LRR 14.  
FT REPEAT 419 444 LRR 15.  
FT REPEAT 446 469 LRR 16.  
FT REPEAT 491 513 LRR 17.  
FT REPEAT 514 537 LRR 18.  
FT REPEAT 539 562 LRR 19.  
FT REPEAT 615 638 LRR 20.  
FT REPEAT 640 663 LRR 21.  
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 782 AA; 88561 MW; 311460B2D4527917 CRC64;

Query Match 7.5%; Score 279; DB 1; Length 782;  
Best Local Similarity 24.8%; Pred. No. 1.9e-12;  
Matches 145; Conservative 94; Mismatches 233; Indels 112; Gaps 23;

QY 47 YRRATVDCNDLELT-AVRRPALRAGQOTLLQSNSTVRRQDSGLGLANTLELDLSONSF 105  
DB 5 HRRQRFVDIGDSLTLTKLTHAL-----SSVQNFPSDAIKILNLEEDLSNNRL 54  
QY 106 SPARDCEFHLPOLLSTLHLEENOLTRLEDHSFAG--LASLOELYLNHNOLYRIAPAFSG 163  
DB 55 RNVDPMSFHFPLSLKRVHLODWTIEIHNGTQGDHRLDTEVFEFNSVRYNQHTFAD 114  
QY 164 LSNLRHLNLSNLLRAIDSRWEMLPNLEILMIGKKVDAIIDMNRPLANTSLVLAQM 223  
DB 115 LLOLEIHLDNKRIEISLERAPNKLKRLMLKGNKIKTIATVEYTONPLEDLDLAVN 174  
QY 224 NIREISDVALEGQSLSELSFYDN-----QLARPRALRQVPL-----KFLDK 270  
DB 175 SISLDFNIFDOVSGIG--MFHVNSHNKLINLVAVSVFFQDTGLGIGIQNIKIVLDSIF 232  
QY 271 NPLQRYPGDPFANM-LHLKELGINMEELVSIDKFALVNLPELTKIDITNNP--RLSFIH 327  
DB 233 NUTTSVAKQFFRPVELSLMQLYLGH-NKLLNMTKDLFGMMPHLQVLDLSHNSIYELDF-- 289  
QY 328 PRAFHHLPMQETLMLNNALSLHOOYVESLPNLOEVGLHGNPIRC--DCVIRMANATGT 385  
DB 290 -DTFRNTRKQLQWLDIYSHNRISIPNDLFFGLNLRIVDESHNRRLRSLPNLFR---ETG- 344  
QY 386 RVRFIEPOSTLCAEPDQRLPLRVREVPFREMDDHCLPLSPSPSPSLOVAGSGSNVLLHC 445  
DB 345 -----LERLDVSHNLGK-----LPL-----TSLSLASQYL----- 371  
QY 446 RALAEPEPEIYVTPAGLRLTPAHAGR--RYRVPEEGTELELRVTAEEAGLYTCV---AQ 500  
DB 372 -----SELDLSNKSISL-----SHGQGLARFKCLSMLDSTYRKLOIDAGTKGIPRLAS 422  
QY 501 NLVGADITVSVVGRALLQPGRDEGQGLELRVQETHPYHILLSWV---TPPNTVSTNLT 557  
DB 423 LNLGNSQLTLEINGLSF-----QGLEYTLILANDVNSLSQVPAISPALSTPNLSTLSLA 474  
QY 558 MSS-----ASSLR--CGGATALARLRPGHSHNITRLL 588  
DB 475 FNSLPTVALEVAGNISLRYNLINDYNDLSAVPIVYHSLTELRHL 518

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 31, 2002, 14:36:13 ; Search time 34.24 Seconds

(without alignments)  
3602.377 Million cell updates/sec

Title: us-09-905-088a-245

Perfect score: 3732

Sequence: 1 MRLVAPLLAWAGATATV.....RKLPSSSEGTLLPPLSQNS 713

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1695	45.4	730	4	09P231
2	1687	45.2	716	11	061809
3	1656.5	44.4	705	4	043377
4	1656.5	44.4	708	4	09H3W5
5	1649.5	44.2	707	11	09ESY6
6	1648.5	44.2	707	11	P97860
7	1625.5	43.6	718	13	073675
8	875.5	23.5	431	4	09H8V1
9	755.5	20.2	273	4	09N0U4
10	494.5	13.3	614	11	09D1T0
11	492.5	13.2	614	4	096FES
12	490	13.1	614	6	09N008
13	468.5	12.6	606	4	09B820
14	442	11.8	1091	11	P70193
15	438.5	11.7	640	4	09HC22
16	432.5	11.6	540	5	09V053

17	431	11.5	1094	4	09BYB8	09byb8 homo sapien
18	430	11.5	1093	4	096JAI	096jai homo sapien
19	427.5	11.5	532	5	096B71	096b71 drosophila
20	420.5	11.3	628	4	09B7N0	09b7n0 homo sapien
21	407	10.9	733	5	024250	024250 drosophila
22	407	10.9	737	5	09VU51	09vu51 drosophila
23	406.5	10.9	719	4	096NT6	096nt6 homo sapien
24	401.5	10.8	737	5	0965M3	0965m3 caenorhabdi
25	401.5	10.8	881	5	0965M2	0965m2 caenorhabdi
26	401.5	10.8	1447	5	016779	016779 caenorhabdi
27	399.5	10.7	811	4	075139	075139 homo sapien
28	399	10.7	78	11	061974	061974 mus musculu
29	398.5	10.7	789	6	09BE71	09be71 macaca fasc
30	395.5	10.6	673	11	09CZT5	09czt5 mus musculu
31	392.5	10.5	792	4	09UDT7	09udt7 homo sapien
32	388.5	10.4	832	4	09ULH4	09ulh4 homo sapien
33	386	10.3	492	11	099K76	099k76 mus musculu
34	386	10.3	653	4	09HBW1	09hbw1 homo sapien
35	385	10.3	788	11	09CYK3	09cyk3 mus musculu
36	378	10.1	1065	4	096A85	096a85 homo sapien
37	377	10.1	1065	4	094898	094898 homo sapien
38	374.5	10.0	700	4	09P244	09p244 homo sapien
39	368	9.9	907	11	09Z1P4	09z1p4 mus musculu
40	361.5	9.7	542	5	09N4G6	09n4g6 caenorhabdi
41	361.5	9.7	809	11	09DBY4	09db4 mus musculu
42	361	9.7	1173	5	09V7J8	09v7j8 drosophila
43	360.5	9.7	1531	11	088279	088279 rattus norv
44	360	9.6	473	11	099P18	099p18 mus musculu
45	360	9.6	907	4	075473	075473 homo sapien

## ALIGNMENTS

RESULT 1  
ID 09P231 PRELIMINARY: PRT: 730 AA.  
AC 09P231  
DI 01-OCT-2000 (TREMBLrel. 15, Created)  
DI 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DI 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE KIAA1497 PROTEIN (FRAGMENT).  
GN KIAA1497.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20277482; PubMed-10819331;  
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Chara O.,  
RT "Prediction of the coding sequences of unidentified human  
RT genes. XVII. The complete sequences of 100 new cDNA clones from brain  
RT which code for large proteins in vitro."  
RL DNA Res. 7:143-150(2000).  
DR EMBL, AB040930; BA96021.1;  
DR InterPro: IPR003961; FN\_III.  
DR InterPro: IPR003598; Iq\_c2.  
DR InterPro: IPR003006; Iq\_MHC.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR.  
DR InterPro: IPR000372; LRR\_Cterm.  
DR InterPro: IPR003592; LRR\_Out.  
DR InterPro: IPR003591; LRR\_Typ.  
DR Pfam: PF00047; Ig\_1.  
DR Pfam: PF00560; LRR\_9.  
DR Pfam: PF01463; LRRCT\_1.  
DR Pfam: PF01462; LRRNT\_1.  
DR SMART: SK00060; FN3\_1.  
DR SMART: SK00408; IGC2\_1.  
DR SMART: SK00370; LRR\_4.  
DR SMART: SK00082; LRRCT\_1.  
DR SMART: SK00013; LRRNT\_1.

```

OY      619  --TSCHRALGDRPGCLAIATLAAVL 641
Db      626 HETSTALAAVWGSFPAVISIAIAYI 652

RESULT      3
ID          043377    PRELIMINARY; PRT; 705 AA.
AC          043377;
DT          01-JUN-1998 (TREMBLrel. 06, Created)
DT          01-JUN-1998 (TREMBLrel. 06, last sequence update)
DE          01-DNC-2001 (TREMBLrel. 19, last annotation update)
DE          PROBABLE LEUCINE-RICH REPEAT PROTEIN.
OS          KGI18D07.1.
OC          Homo sapiens (Human).
OC          Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN          NCBI_TaxId=9606;
RP          SEQUENCE FROM N.A.
RS          TISSUE=FIBROBLAST;
RA          Murray J., Langston Y., Clarke C.;
RL          Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
DR          EMBL: AC004142; AAC02752.1; -.
DR          InterPro: IPR0033961; FN_111.
DR          InterPro: IPR003398; Ig_C2.
DR          InterPro: IPR003006; Ig_MHC.
DR          InterPro: IPR001611; LRR.
DR          InterPro: IPR000483; LRR_Cterm.
DR          InterPro: IPR000372; LRR_Nterm.
DR          InterPro: IPR003592; LRR_Out.
DR          InterPro: IPR003581; LRR_Lyp.
DR          Pfam: PF00041; fn3; 1.
DR          Pfam: PF00047; Ig; 1.
DR          Pfam: PF00560; LRR; 9.
DR          Pfam: PF01463; LRRCy; 1.
DR          Pfam: PF01462; LRNT; 1.
DR          PRINTS: PR00019; LEURICRPT.
DR          SMART: SM00408; IGC2; 1.
DR          SMART: SM00370; LRR; 4.
DR          SMART: SM00082; LRRCy; 1.
DR          SMART: SM00013; LRNT; 1.
DR          SMART: SM00369; LRR_Typ; 1.
KW          Immunoglobulin domain.
SQ          SEQUENCE 705 AA; 79049 MW; 9BA6066693F4BC1A CRC64;

Query Match      44.4%; Score 1656.5; DB 4; Length 705;
Best Local Similarity 47.2%; Pred. No. 6e-118;
Matches 333; Conservative 109; Mismatches 217; Indels 47; Gaps 5

OY      17 TATVEVPMWHVCPPQCAQIRPWTPFRSSYEATVTDCNDLFLFAVPALPAGTGITLL 76
Db      14 TLTVAVADKKVDCPRLCTGEIRPWTPTSRIWEASIVDCNDGLTLFPARLPANTOITLL 73
OY      77 ONSNIVRPDQSLGILIANLTELDLSONSFSDARDCEFHALLPOLLSIHLENDOTLRLEDHS 136
Db      74 QTNNAKTIETS -TDPRPVNLGTLDLSONNISVTYNIVKKMPOLLASYILENKITELPEKC 132
OY      137 FAGLASLOELEYNNHQLRAPRASGLSNLRHLHNSULRALAISRWEMLPNEILMT 196
Db      133 LSELNDELBYLNHMLSTISPGATIGLHNLRLHNSNRLOMINSKWDALPNIEILMT 192
OY      197 GGNKVDAIDLMMNFRLPLANLSVLGMMILREISDALLEGIOGLESTSPFYDNOLAVPRRA 256
Db      193 GENPIIRIKDMFKLINLRSLVIANGINLTETPDNALVGLENBESISTEDNRLLIVPHYA 252
OY      257 LEQVGLKFLDLINKPDLORVGPDFANNMLHKELGLNNNEELVSIDKFALVNLPELTKD 316
Db      253 LQKYVNLFELDLINKPKINRKGRDPFSNMHLKEKLGINNMPELTSIDSLADVNLPLRKIE 312
OY      317 ITNPRLFSEIFRPRAHHHPOMETTMLNNNALSAHQOVVESIPNQOEVLGSGNPTRCQC 376

```

Db	Accession	Source	Length	Score	E-value	Bit Score	Identities	Positives	Gaps	Matches	Mismatches	Indels	Frameshifts	Insertions	Deletions	Repeats	Other	Notes
Db	313	ATNNPLSTYIHPHAFERLKLSSLMNLSNALSALYHGITESLPNIKEIISHSNFRIDCV	372															
Qy	377	IRNANAATGTRVRETFIEFOSTLCAEPDLOLRPYREVFPREMTDCLPLISPRSPSLQVA	436															
Db	373	IRMNNAKTKTIREMEPDSLFCDVDFEFGQNVQVHFHDMDEICLPLAPSPSPNINVE	432															
Qy	437	SGSWSVLHCRALAEPEPELTYWTPRGLKLTFAHAGRRRYRYPESTLELRRYLAEGLYT	496															
Db	433	AGSYVSEHCRARAEPOPELTYWTPRGLKLTFAHAGRRRYRYPESTLELRRYLAEGLYT	492															
Qy	497	CVANQNVAGDFTSVYVVRALLOPGRDEGGLELRVQETHPRYHILLSWTPPNRYSTNL	556															
Db	493	CITNVLGADLKSVMIKDGSPPO--DNNSGLWIKTDIOANSVLYSWASSKILKSSV	549															
Qy	557	TWSSASSLGGQATALARLPFGTHSYNITRLLOATEYACLOVAFADAHQOLACVARTK	616															
Db	550	KMTAFYKTENSHAQASRIPSDVKYNNLTNLHPNSTEYKICIDIFTIYQKRKKCVWYTK	609															
Qy	617	-----EATSCRALGDRPGLIATLAILALVLLAAGLANHIGTQPRKGVGRR--	663															
Db	610	GLHPDQKEYEKNNNTTLMACLGGLIGIIGYACILSCL-----SPENCDGSHS	657															
Qy	664	-----PLPPANAFQWQASPSRVYVASAPLYLPMN	691															
Db	658	YVRNLTQKPTPALGELYPLINLMEAGKEKSTLKVATYIGLPTN	703															
RESULT	4																	
Q9H3W5																		
ID	Q9H3W5	PRELIMINARY;	PRY;	708	AA.													
AC	Q9H3W5	01-MAR-2001 (TREMBLrel. 16, Created)																
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)																	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)																	
DE	HYPOTHEICAL 79.4 KDA PROTEIN (NEURONAL LECUCINE-RICH REPEAT																	
GN	DKEPZP61K2424 OR NLRR-3.																	
OS	Homo sapiens (Human).																	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.																	
OX	NCBI_TaxID=9606;																	
RN	[1]																	
RP	SEQUENCE FROM N.A.																	
RC	TISSUE=AMYGDLA;																	
RA	Blocker H., Boecker M., Brandt P., Mewes H.W., Weill B., Wiemann S.;																	
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.																	
RN	[2]																	
RP	SEQUENCE FROM N.A.																	
RA	Hamano S., Inuzuka H., Morohashi A., Ohira M., Nakagawara A.;																	
RT	"Human neuronal leucine-rich repeat protein-3(NLRR-3).";																	
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.																	
DR	EMBL; AL42092; CAC09450.1; -																	
DR	EMBL; AB060967; BAB47184.1; -																	

```

Db 483 IIGITPREGGLTICATFNVLGADLKSIMIKVGGFVPO---DNGSLINIKIRDIRANSVLY 539
QY 544 SVTPPNTVSTNLTWSSASSLRQGATATALRLPRGTHSYNITRLLOATEYMACLOYAFAD 603
Db 540 SWKANSKILKSSVKMTAFVKTEDSOQAASARIPSDVKVYNLTLLKSTEEKICIDIPITY 599
QY 604 ATTQALACVART-----KEATSCH----RALGDRPGLAIATLAVLLAAGLAHLGT 652
Db 600 QKSRQCVVNTTKLSLEHDKENKSHTEVACVGLIGIVMCL-----FCVSGOEGN 653
QY 653 GQPRKGVGRRLPPAMAFMGNSAPSVRYSAFLVLPMPNGRLPSSSE 701
Db 654 GENESYTVNHCHKPTLAF-----SELYPPLINLMESSKEKPASLE 694

```

```

RESULT 6
P97860 PRELIMINARY; PRT; 707 AA.
AC P97860.
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LECICINE-RICH REPEAT PROTEIN PRECURSOR (FRAGMENT).
GN LRRN3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96233817; PubMed=9011764;
RA Taniuchi H., Tohyama M., Takagi T.;
RT "Cloning and expression of a novel gene for a protein with leucine-
RL rich repeats in the developing mouse nervous system.";
RL Brain Res. Mol. Brain Res. 36:45-52(1996).
DR EML; D49802; BAA08622.1; -.
DR MGD; MGI:106036; Lrrn3.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_tyr.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; fn3; 1.
DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_tyr; 1.
KM Immunoglobulin domain; Signal.
FT SIGNAL 1 20
FT CHAIN 21 >707 POTENTIAL.
FT NON_TER 707 707
SQ SEQUENCE 707 AA; 79156 MW; FAETC1573DDDL65B CRC64;

```

```

QY 17 TATPVVPMHPCPPQCCOIRPMPTSSRYREATTVCNDLFLAVPALPAGQTILL 76
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 17 TTLVQALDKKDCFCOLCEIRPMFTPRSTYMEASTVDCNDGLNLPARLPADTOILL 76

```

Query Match 44.2%; Score 1648.5; DB 11; Length 707;  
Best Local Similarity 47.7%; Pred. No. 2.5e-117;  
Matches 328; Conservative 111; Mismatches 225; Indels 23; Gaps 5;

```

QY 77 QNSNIYRVDSSELGYLANITELDLSONSFEDARDCDFHALPOLLSTLSENOTRLDEHS 136
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 77 QTNINRIIEHS-IDFPVNLGIDLSQNNLSVTNINQKSOQLSYLEENKITELEK 135
QY 137 FAGLASLOEYLYNNOLYRIAPAFSGLSNLRHLHNSNLRRAIDSRWFEMLPNLEITMT 196
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 136 LVGLSNLOEYLYNNHLSLSTSPGAFIGLHNLRLHLSNRLQMINSQWFALPNLEITML 195
QY 197 GGNKVDAILDMNRPRLANRSYLACMNLREISDYALEGQSLESLSFYNOIARPPRA 256
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 196 GDNPIIRIKDMNQPVLKRSLSVIAGINTEIPDDALAGENLESISFYDNRLSKYPOVA 255
QY 257 LEQVPLKFLDLNKNLQVRGDFANMLHLKELGNNMEELISIRKFAVLNLPETLKD 316
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 256 LQKAVNLKFLDLNKNINRTRRDEFSNMLHLKELGINNPELVSIDSLAVNDLPDLKTE 315
QY 317 TNNPRLSFTHPPAFHLLPOMETIMLNNAALHQOYVESLPNLOEVLGNPRICDCV 376
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 316 ATNNPRLSYTHPPAFHLLPRLKESIMLTNNAALSYHGTEISLPNLEKISHSNPRICDCV 375
QY 377 IRANANTGRVREIEPQSLCAEPDQLRVPREVPFRENTDCLPLISPRSPSLQVA 436
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 376 IRWINNKNINIREMPSLFCVDPPEFQONKQVHRDMMETCLPLIAESEPDLDER 435
QY 437 SGEWVILHCRALAEPEPEIYWTAGLRTPAHAGRRYRVPGETLELRVTAEEAGLYT 496
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 436 ADSYVSLHCRATAEPPEIYMTIPSSKLLPNTMRKFPYHSGTIEITPKEGLYT 495
QY 497 CVAQNLVGADTVSVVGRALLQGRDEGQGLERVOETHPHYLLSWTPPNTVSTNL 556
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 496 CIATNLVGADLKSIMIKVGGSVPO---DNGSLINIKIRDIRANSVLYSVKASKILKSRY 552
QY 557 TWSSASSLRQGATATALRLPRGTHSYNITRLLOATEYMACLOYAFADARTQALCVARTK 616
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 553 KWTGEVKTEDSHAAQASARIPSDVKVYNLTLLKSTEEKICIDIPITYOKSRQCVVNTTK 612
QY 617 EATSCHRALGDR-----PGLAITLAVLLAAGLAHLGTGQPRKGVGRRLPPA 668
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 613 SLEHDKKEYGNHITVYVACVGLIGIVMCLF---SCVSQBSGSGEHSYAVNHCHKPA 669
QY 669 WAFMGNSAPSVRYSAFLVLPMPNGRK 695
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 670 LAF-----SELYPPLINLMESSKE 688

```

```

RESULT 7
ID 073675 PRELIMINARY; PRT; 718 AA.
AC 073675;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NEURONAL LECICINE-RICH REPEAT PROTEIN.
GN XNLR-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesodactylia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99070063; PubMed=9852961;
RA Hayata T., Uochi T., Asashima M.;
RT "Molecular cloning of XNLR-1, a Xenopus homolog of mouse neuronal
RT leucine-rich repeat protein expressed in the developing xenopus
RL nervous system.";
RL Gene 221:159-166(1998).
DR EML; AB014462; BAA28530.1; -.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.

```

DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE PLAC1006239 PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugeno S., Ishibashi T., Fujimori K.,  
 RA Tanai H., Kimata M., Watanabe M., Hirooka S., Ishii S., Kawai Y.,  
 RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,  
 RA Masuko Y., Kanehori K.;  
 RT "NEO human cDNA sequencing project";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK001991; BAA92025.1; -;  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003592; LRR\_out.  
 DR InterPro; IPR003591; LRR\_typ.  
 DR Pfam; PF00560; LRR; 6.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR PRINTS; PRO0019; LEURICHRPT.  
 DR SMART; SM00370; LRR; 2.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 1.  
 DR NON\_TER 273  
 FT SEQUENCE 273 AA; 30832 MW; B59F070A6D7D11FF CRC64;

Query Match 20.2%; Score 755.5; DB 4; Length 273;  
 Best Local Similarity 58.8%; Pred. No. 1e-45; Mismatches 66; Indels 1; Gaps 1;  
 Matches 151; Conservative 39; Mismatches 66; Indels 1; Gaps 1;  
 QY 17 TATVAVPVHVPPOCAQIRPWYTPRSSYREATTDCNDLFTAVPAPAGTQTL 76  
 DB 17 TLVAVADKKVDCPRLCTGCEIRPWFPRSIWEASVDCSDGLLTFPRLPANTOILL 76  
 QY 77 GNSIVRVDQSELGYANTTELDLSONSFSDACDFHALPOLLSHLEENOTRL 136  
 DB 77 GNNIAKLEYS-TDEPVNLGDLSCNNLSVTYNIVKMKAPOLLVYLEENKLTPEK 135  
 QY 137 FAGLASLQELYNHNOQYRIAPRAFSGLSNLRHLNSNLRAIDSRWEMPNLEI 196  
 DB 136 LSELNQLQELYNHNLSTISPGAFIGLNLRLHLNSNRLQWINSKWDALPNLEI 195  
 QY 197 GKNKDAILDNFRPLANRLSLVLAGMNLREISDYALEGLQSLSTSFYDNO 256  
 DB 196 GNPPIIRIKDMFKPLINRLSVIAGINLTLPDNLVLENLSTSFYDNLRIKYP 255  
 QY 257 LEQVPGKFLDKNKP 273  
 DB 256 LQKVNLKFLDKNKP 272  
 RESULT 10  
 Q9D1T0 PRELIMINARY; PRT; 614 AA.  
 AC Q9D1T0;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE ADULT MALE TESTIS CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
 DE CLONE:4930471K13, FULL INSERT SEQUENCE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=TESTIS;  
 RX MEDLINE=21085660; Pubmed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bouo H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Mikaldo I., Pesole G., Quackenbush J.,  
 RA Schirni L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombauts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK027262; BAB32403.1; -;  
 DR HSSP; P23945; 1XUN.  
 DR InterPro; IPR003599; I9.  
 DR InterPro; IPR003598; I9\_C2.  
 DR InterPro; IPR003600; I9\_Like.  
 DR InterPro; IPR003006; I9\_MHC.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR Pfam; PF00047; I9; 1.  
 DR Pfam; PF00560; LRR; 9.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR SMART; SM00409; I9; 1.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00410; I9\_Like; 1.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 9.  
 KW Immunoglobulin domain.  
 SO SEQUENCE 614 AA; 69100 MW; 41CFP40C21335681 CRC64;

Query Match 13.3%; Score 494.5; DB 11; Length 614;  
 Best Local Similarity 25.6%; Pred. No. 2.9e-29;  
 Matches 173; Conservative 97; Mismatches 251; Indels 155; Gaps 18;  
 QY 1 MRLVAPLLAN-----VAGTAVPVVHVPPOCAQIRPWYTPRSSYREA 50  
 DB 6 MRSMPSPLACQPIILLVIGSVLSGSAF-----GCPPRECS-----AGD 46  
 QY 51 TVVDCNDLFTAVPAPAGTQTLTILQNSIVRVDQSELGYANTTELDLSONSFSDAD 110  
 DB 47 RAVLCHRRKRFVAVPBGIPETRLDLGNRIKRLTNDDEFASPHLEELNENIVASVP 106  
 QY 111 CDFH-----ALPOLLSHLEENOTRLREDSFAGLASLOEL 146  
 DB 107 GAFNNLNFNRTIGLSNRLKLPGLVFTGLSNLTIDTIDISENKIVILLDYMFO 166  
 QY 147 YNHNQYRIAPRAFSGLSNLRHLNSNLRAIDSRWEMPNLEIIMGNKVDALID 206  
 DB 167 EVGDNDLVYISHRASGSLSTEQLTLEKCNLTISFTEALSHLGLIVLRHLNINAI 226  
 QY 207 MNERPLANRLSLVLAGM-NLREISDYALEGLQSLSTSFYDNOALVRRALQVPGKLF 265  
 DB 227 YSFKRLRYKLYEISHWPYLDTPNCLYGL-NLSTLSTHCNLTAVPYLAVRHLYLRF 285  
 QY 266 IDLNKNPQVRGDPGFANMLHLKELGLNNMELVSIIDKFAVLNLPDLRYKLDITNNPRISF 325  
 DB 286 LNLSTNPIGTIF---GSMTH-----ELNQLQELQV-----GGQLAV 319



```

QY 1 MLLVAPLLIAW-----VAGATATVPVWVWPCPOCAQOIRPWYTPRSSYREA 50
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 VRSPSPLLACWOPILLVGLVSGSAT-----GCPRECECS-----AOD 46
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 51 TTYDCNDLFLTANPPALPACTOILLQSNSTVAVDSEGLYANLVELDSONSPSDARD 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 47 RALVCHRRKRPVAAVEGIPETRIIDLGKNNIKIKLNODEFASFPHELEENINIVASVP 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 CDFH-----ALPOLLSHLENOLTRLEDHSHFALASLOEL 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 GAFNNIFNLRTGLBRNRLKILPGVFTGISNLTIKDISENKIVILLDYFODLYNKLKL 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 YLNHNOLYIAPRAFSGLSNLRLHLNSMLRAIDSRWEMLPNLEILMIGKNKYDAIID 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 EVGDNDLYVISHRAFSGLSNLEQLEKCKLSTIPREALSHLGLVLRHLNINAIRD 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 MNPRLANRSLVLGAM-NLRISDYALBGLSLESLSFYDNOANVPRLAEQVGLKF 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 YSFKRLYRLKVLKLEISHWPYLDITWTPNCLVGL-NLTSLSTHCHMTAVPYLAVERHLVYLR 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 LDLNKPNLQVGRGDPANMLHLKEGLGNMNEELVSDKALVNLPELTIKDITNNRSLF 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 LNLSTNPITIE-----GSMH-----ELRLQETOLV-----GGOLAM 319
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 326 IHPRAFHLPQMETMLNNAALSALHQOVESLPNLEQVGLHGNPITRCDCVIRMANATGT 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 VEPYAFARGNLYRLVNLVSGNQLTTLEESVFHSGNLETLIDSNPLACCCORLIMVRRRM 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 RYRFTIEPOSTLCAEPDLORLVREVPREMTDHCPLISPSNF-----PPSLQ 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 RLNFNRQOFT-CATPEFVGGKEKDFP-----DVLIPNYFTCRARIRDRKAQOYF 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 435 VASGESMVHCRALAEPEPEIYWTVPAGLRTPAHAGRYRYVPEGTLELRTATAEAGL 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 430 VDEGHTVQVCAADGDPRAIILMSLR-KHLYSAKSNGLTYFPDSTLEKRAQVODNGT 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 495 YTCVAVNLVADT-----KTVSVVGRALLQGRDEGQLELRVOE 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 YLCIANANGNDSMPAHLHVRSYSPDMPHQPKTFATFIN-----QPG--EGEANGSTRATV 542
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 536 TTPYHI 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 543 PPPFDI 548
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
Q9BZ20 PRELIMINARY: PRT: 606 AA.
AC Q9BZ20 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE BA438B23.1 (NEURONAL LECINE-RICH REPEAT PROTEIN) (CDNA FLJ1810 FTS,
CLONE NT2121009289, WEAKLY SIMILAR TO CARBOXYPEPTIDASE N 83 KDA
CHAIN).
GN BA438B23.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ishidashi T., Kanehori K., Yosida M., Matanabe S., Ishida S., Ono Y.,
RA Houta T., Hirakawa S., Murekawa K., Takiguchi S., Kusano J.,
RA Matanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,

```

```

RA Magatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.,
RA "NEO" human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL353746; CAC22713.1;
DR EMBL: AK056372; BAB71167.1;
DR InterPro: IPR003599; Ig_
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF00560; LRR_11.
DR Pfam: PF01463; LRRCT; 1.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00409; IG_1.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 10.
KW Immunoglobulin domain.
SO SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;

Query Match 12.6%; Score 468.5; DB 4; Length 606;
Best Local Similarity 27.0%; Pred. No. 2,7e-27;
Matches 140; Conservative 84; Mismatches 221; Indels 73; Gaps 9;

QY 27 VCPPOCAQOIRPWTPRSSYREATTVCNDLFLAVPALPAGTQTLILQSNSTVAVDQ 86
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 ICGPACCECS-----AQNKSVCCHRRRLIAIPGIPLETFTK----- 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 SELGIANTLELDLSONSFSARDCDFHALPQLLSHLENOLTRLEDHSHFALASLOEL 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 -----LDLSKNRLKSVNPEEFISYPLLEETLSDNTIANVEPAFNLLFNLRL 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 YLNHNOLYIAPRAFSGLSNLRLHLNSMLRAIDSRWEMLPNLEILMIGKNKYDAIID 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 RLKGNRLKLVPGVFTGLSNLTIKDISENKIYIILDYMODLHNLKSLVGDNDLYISH 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 MNPRLANRSLVLGAM-NLRISDYALBGLSLESLSFYDNOANVPRLAEQVGLKF 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 RAFSGLSLEQLTLKCNLTAVPTALSHLRSLISLHKLNINMNPVYAFKRLFKHL 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 DLNKNPDLQVGRGDPANMLHLKEGLGNMNEELVSDKALVNLPELTIKDITNNP----- 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 EIDVYPLDMMPANSLYGLNLTSLSVTN-TNLSVPEFLAKFLVYLTNLTSLNPNSTIE 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 322 -----RLSFTHPRAFHLPQMETMLNNAALSALHQOVESLPNLOE 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 290 AGMFSDLIRLOELHVGQALRTIEPHSFGGLRFLVIANSONMLLEENYSSPRALEY 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 364 VGLHGNPITRCDCVIRMANATGTRVRFIEPQSTLCAEPDLORLVREVPREMTDHCPL 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 LSINNPLACDCRLIMILOROPTLOF-GGQOPMCAGPD-----TIREPSFKDFHSTALS 403
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 424 I-----SPRSFPSLQ-----VASGESMVHCRALAEPEPEIYWTVPAGLRTPAHAGRYRYV 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 404 YTCCKPKPIREKKLOHLVDSQTVQLDSCAGDPQVPSWTPRRRTTKSNGRA-TV 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 477 YPEGTLELRVYAEAGLYTCAONLVGADRTVSVVV 514
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 463 LGDGLLEIRFQDDSGMYVCVCLASNAANDPTTASTLV 500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 14
P70193 PRELIMINARY: PRT: 1091 AA.
ID P70193

```



